Synthese Library 373

Alessandro Torza *Editor*

Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language

Springer

Synthese Library

Studies in Epistemology, Logic, Methodology,
and Philosophy of Science

Volume 373

Editor-in-Chief
Otávio Bueno, University of Miami, Department of Philosophy, USA

Editors
Dirk van Dalen, University of Utrecht, The Netherlands
Theo A.F. Kuipers, University of Groningen, The Netherlands
Teddy Seidenfeld, Carnegie Mellon University, Pittsburgh, PA, USA
Patrick Suppes, Stanford University, CA, USA
Jan Wolenski, Jagiellonian University, Kraków, Poland

More information about this series at http://www.springer.com/series/6607

Alessandro Torza
Editor

Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language

Springer

Editor
Alessandro Torza
Instituto de Investigaciones Filosóficas, UNAM
Circuito Mario de la Cueva, Ciudad Universitaria
Del. Coyoacán, México D.F., Mexico

Synthese Library
ISBN 978-3-319-18361-9 ISBN 978-3-319-18362-6 (eBook)
DOI 10.1007/978-3-319-18362-6

Library of Congress Control Number: 2015945970

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Contents

Contributors

Ken Akiba Department of Philosophy, Virginia Commonwealth University, Richmond, VA, USA

G. Aldo Antonelli Philosophy Department, UC Davis, Davis, CA, USA

Axel Arturo Barceló Aspeitia Instituto de Investigaciones Filosóficas, UNAM, Circuito Mario de la Cueva, Ciudad Universitaria, Del. Coyoacán, México D.F., Mexico

Sharon Berry School of Philosophy, Australian National University, Canberra, ACT, Australia

Francesco Berto Department of Philosophy and Institute for Logic, Language and Computation (ILLC), University of Amsterdam, GC Amsterdam, The Netherlands

Solomon Feferman Department of Mathematics, Stanford University, Stanford CA, USA

Mario Gómez-Torrente Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, México D.F., Mexico

Andrea Iacona Centro di Logica, Linguaggio e Cognizione, Dipartimento di Filosofia e Scienze dell'Educazione, Università di Torino, Torino, Italy

Bjørn Jespersen LOGOS, Departament de lògica, història i filosofia de la ciència, Universitat de Barcelona, Barcelona, Spain

Arnold Koslow Philosophy Program, The Graduate Center, CUNY, New York, NY, USA

Gregory Lavers Department of Philosophy, Concordia University, Montreal, QC, Canada

David Liebesman Department of Philosophy, The University of Calgary, Calgary, AB, Canada

Friederike Moltmann Institut d'Histoire et de Philosophie de Sciences et Techniques (IHPST), Université Paris 1, Paris, France

Sun-Joo Shin Department of Philosophy, Yale University, New Haven, CT, USA

J.P. Studd Lady Margaret Hall, University of Oxford, Norham Gardens, Oxford, UK

Alessandro Torza Instituto de Investigaciones Filosóficas, UNAM, Circuito Mario de la Cueva, Ciudad Universitaria, Del. Coyoacán, México D.F., Mexico

Tero Tulenheimo STL-CNRS and Department of Philosophy, University of Lille 3, Domaine Universitaire du "Pont de Bois", Villeneuve d'Ascq, France

Jason Turner Department of Philosophy, Saint Louis University, St. Louis, MO, USA

Alberto Voltolini Dipartimento di Filosofia e Scienze dell'Educazione, Universitá degli Studi di Torino, Torino, Italy

Elia Zardini FCT Research Fellow, LanCog, Language, Mind and Cognition Research Group, Centro de filosofia, Universidade de Lisboa, Lisboa, Portugal

Chapter 1
Introduction

Alessandro Torza

Abstract This introductory chapter provides a summary of the contributions to the volume, as well as some critical remarks.

The development of formalized quantificational languages is one of the most groundbreaking events to ever take place in the history of philosophy. Since the work of Frege, quantifiers have played a crucial role in the introduction of new philosophical concepts and paradigms, as well as in the analysis and clarification of philosophical arguments. The centrality of the predicate calculus as a philosophical *lingua franca* has spurred a host of studies on the logic of quantifiers, their interaction with other logical operators, the role they play in assessing ontological debates, their significance in the foundations of mathematics, their usefulness in the analysis of expressions not (overtly) quantificational, the coherence of absolute generality—and related topics. As a result, we can now speak of a *philosophy of quantifiers* which spans across multiple areas such as logic, philosophy of language, metaphysics, epistemology and even the history of philosophy. Nevertheless, a collection of essays on the philosophy of quantifiers had been conspicuously missing. That absence is the motivation behind the present volume, which aims to cover a number of contemporary issues surrounding the nature of quantification. The essays in this collection are grouped according to five main themes: *logical constants*; *the semantics of natural language*; *the Carnap-Quine legacy*; *metaphysics and ontology*; *logical systems*.

Following Alfred Tarski, the relation of logical consequence can be characterized semantically: ϕ is a logical consequence of a set of sentences Γ just in case ϕ is true in every model of the language which makes true every member of Γ. The range of the possible models is defined by varying the interpretation of the non-logical constants while keeping the interpretation of the logical constants fixed. Hence, Tarski's semantic notion of logical consequence presupposes a demarcation

A. Torza (✉)
Instituto de Investigaciones Filosóficas, UNAM, Circuito Mario de la Cueva, Ciudad Universitaria, Del. Coyoacán, México D.F. 04510, Mexico
e-mail: atorza@me.com

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_1

between *logical* and *non-logical constants*. Tradition has it that the logical constants should include at least truth-functional connectives and quantifiers. In her essay 'Quantifiers are Logical Constants, but Only Ambiguously', Sun-Joo Shin argues that there is not "much in common between sentential connectives (which are taken to be logical constants without any controversy) and quantifiers". For connectives have constant interpretation across models (as defined by the respective truth functions), whereas the interpretation of a quantifier changes with the domains. The existential quantifier, for instance, can be interpreted as the set of the non-empty subsets of a domain, and such nonempty sets will be quite different on a domain of natural numbers as opposed to a domain of people. So, if interpretation is extension, quantifiers do not count as logical in a Tarskian setting. Nevertheless, Shin concedes that there exists a second, model-invariant notion of quantifier interpretation. In its general form, this second notion corresponds to that of a Lindström quantifier, which is a function from models to relations on subsets of that model (Lindström [17]). For instance, the quantifier *All* maps each model D to a relation All_D such that, for all subsets $A, B \subseteq D$, $All_D(A, B)$ iff $A \subseteq B$. In this sense it is quite clear that quantifiers do have constant interpretation, whereas quantifier extensions, which are domain-relative, do not.[1]

The issue of providing a characterization of logical constants has proven particularly unwieldy, as amply discussed in Gómez-Torrente [11]. In his 'Which Quantifiers are Logical? A Combined Semantical and Inferential Criterion', Solomon Feferman puts forward a way to draw the logical vs. non-logical distinction in the special case of quantifiers. Feferman's proposal, which is a modification of one by Zucker [33], is that a (Lindström) quantifier is logical only if there exists an inferential condition that uniquely characterizes its interpretation in every model. For instance, the first-order universal quantifier meets the condition insofar as its extension in every model (i.e., the set containing the whole domain as its sole member) is captured by the introduction and elimination rules for $\forall$. *Mutatis mutandis* for $\exists$. Feferman's main theorem states that a quantifier meets this criterion just in case it is definable in first-order logic (FOL).[2] It follows from the theorem that many generalized quantifiers (*half of*, *most*, etc.) as well as high-order quantifiers are not logical. As Feferman points out, it is worth realizing that his syntactical-cum-semantical criterion can only provide a necessary condition for logicality. Indeed, there are group-theoretic notions which can be defined in terms of a Lindström quantifier, but which would hardly count as logical. This is an important dialectical point, since most of the objections to the extant characterizations of logicality concern the *sufficiency* side (cf. Gómez-Torrente [11]). So, it appears that the issue of formulating necessary and sufficient logicality conditions remains an open problem, even in the special case of quantifiers.

---

[1] The difference between the two notions of quantifier interpretation is analyzed in Andrea Iacona's essay in this volume. Cf. Feferman's distinction between global and local quantifier (this volume).

[2] Interestingly enough, in his essay Iacona arrives independently at the same conclusion.

Arnold Koslow puts forward a purely structural criterion of logicality in his essay 'Implicit Definitions, Second-Order Quantifiers, and the Robustness of the Logical Operators'. He adopts an approach to logic in which (i) implication is defined via Gentzen-style structural rules, and (ii) constants are defined via filter conditions on the resulting implication structure $(S, \Rightarrow)$, for S a set of propositions. In the case of the universal quantifier, for example, $\forall xPx$ is defined as the weakest element on the implication structure satisfying the filter condition $[X \Rightarrow Pa$, for all $a]$. The criterion put forward by Koslow is that a constant is logical only if it is implicitly defined[3] by a filter condition, and (ii) the filter is non-trivial.[4] Koslow shows that most of the constants that are usually regarded as logical meet the criterion. This approach differs in at least two major ways from Feferman's discussed above. First of all, it is a characterization of logicality which applies to any term, not only quantifiers. Moreover, Koslow shows that second-order quantifiers would count as logical, too.

It was remarked above that the crucial objections to the extant characterizations of logicality concern their sufficiency. Koslow does not exclude the possibility that his criterion could also work as a sufficient logicality condition. If that was the case, his proposal would differ from Feferman's in a third, important way. But notice that in many systems of logic we find the propositional constants *Verum* ($\top$) and *Falsum* ($\bot$). It is reasonable to expect that, under any given logicality criterion, they should both count as logical or both as non-logical. Within Koslow's structuralist approach, we could implicitly define *Verum* as the least object satisfying the filter condition $[X \Rightarrow X]$, i.e., as the weakest element in S that implies itself. And we could implicitly define *Falsum* as the least object satisfying the filter condition $[X \Rightarrow p$, for all $p \in S]$, i.e., as the weakest element in S that implies every proposition. Since the filter associated with *Verum* is trivial (i.e., it always coincides with S), it doesn't meet the logicality criterion. So, *Verum* is not logical. On the other hand, the associated filter of *Falsum* is non-trivial (since its extension is not universal in any structure of at least two elements such that one doesn't entail the other). But if Koslow's criterion were a sufficient logicality condition, *Falsum* would then be logical. As a consequence, it would *not* be the case that *Verum* and *Falsum* are on a par with respect to logicality. So, there seem to be reasons to believe that Koslow's criterion fails to provide a sufficient logicality condition, just like Feferman's and many others on the market.[5]

The second main topic of this volume is the relevance of quantification theory to the *semantics of natural language*. In his essay 'Conjunctive, Disjunctive,

---

[3]Koslow formulates a suitable generalization of implicit definition in the sense of Beth [3, 4].

[4]A filter is trivial when, for every implication structure $(S, \Rightarrow)$, it coincides with S.

[5]It is worth remarking that *Verum* and *Falsum* are defined here as 0-adic operators on an implication structure, that is, as functions from the empty set into S. On the other hand, Koslow defines operators on an implication structure (such as negation and conjunction) as functions from S (or a Cartesian product of S) into S. Therefore, the above argument assumes a generalization of Koslow's notion of operator on an implication structure. One might object to the generalization on the grounds that it violates Frege's distinction between objects and functions. (I owe this last observation to Arnold Koslow.)

Negative Objects and Generalized Quantification', Ken Akiba deals with one central desideratum of linguistics, namely to provide a uniform compositional account of the truth conditions of

1. Socrates is Athenian
2. Somebody is Athenian

The classical Frege-Russell account doesn't meet the desideratum, since in (1) it assigns to 'Socrates' an individual and to 'Athenian' a first-order property, i.e. a function from individuals to truth values; whereas in (2) it assigns to 'somebody' a second-order property, namely a function from first-order properties to truth values. So, in order to evaluate the above sentences, in (1) we apply the semantic value of 'Athenian' to the semantic value of 'Socrates', which returns the value Truth; whereas in (2) we apply the semantic value of 'somebody' to the semantic value of 'Athenian', which also returns the value Truth (since the set of Athenians is non-empty). Montague [20] offered an alternative account which yields the desired uniform compositional account. He achieved this by treating names like quantifier expressions (to be more specific, like generalized quantifiers), hence as having second-order properties as semantic values. Accordingly, the semantic value of 'Socrates' would be the property of being a property had by Socrates. Thus, we can evaluate (1) by applying the semantic value of 'Socrates' to the semantic value of 'Athenian'. Akiba criticizes the Montagovian approach, insofar as (i) it regards names as denoting properties, rather than individuals, and therefore (ii) it appears to be incompatible with a Kripkean view of names as referring to individuals directly. Akiba's solution is that, instead of lifting names to the type of quantifiers (i.e. second order properties), we should lower quantifiers to the type of names and regard them as denoting individuals. Accordingly, there will be the individual *everybody*, which has all and only the properties that every person has, and *somebody*, which has all and only the properties that some person has, etc. We can now interpret (2) in a way which is compositional and uniform with the interpretation of (1), namely we apply the semantic value of 'Athenian' (a first-order property) to the semantic value of 'somebody' (an individual). This is possible as long as we accept an abundant ontology comprising individuals which are quantificational in nature (as well as conjunctive, disjunctive or negative).

The interplay of ontological commitment and the semantics of natural language is also the topic of Friederike Moltmann's essay 'Quantification with Intentional and with Intensional Verbs', where she claims that language provides evidence in favor of non-existent objects (which she refers to as *intentional*). The need for postulating non-existent objects would follow from two kinds of constructs:

Negative existentials: There is a woman John is thinking about who does not exist
Sentences with transitive intentional verbs: John mentioned some woman who does not exist

Moltmann puts forward two key theses about non-existent objects. First of all, she claims that negative existentials can be true only in presence of a relative clause featuring an intentional verb. For instance, 'There is a woman John is thinking about who does not exist' can be true, whereas 'There is a woman who does not exist' can't. However, this claim seems problematic since the former sentence entails the latter and therefore, if the latter is not satisfiable, neither is the former. But the possible truth of the former was supposed to provide evidence in favor of non-existent objects.

Moltmann also claims that nonexistent objects "are not part of the ontology", insofar as they are the result of failed acts of intentionality (cf. McGinn [18]). However, this claim is hard to reconcile with the fact that (i) according to Moltmann, we quantify over non-existents (e.g., in negative existentials), and (ii) ontology is typically regarded as the theory of what *there is*, unrestrictedly. This suggests that Moltmann is presupposing a non-standard notion of ontology, perhaps as a theory of what has being or what exists. (See below Francesco Berto's remarks on such distinctions as part of a defense of Meinongianism.)

'Quantification and Logical Form', by Andrea Iacona, is an attempt at rectifying a common view on what it means to formalize quantified statements. First of all, Iacona distinguishes between first-order *definability* and the weaker condition of first-order *expressibility* of a quantified expression. For instance, *All* is first-order definable insofar as there exists a first-order sentence, $\forall x(Px \to Qx)$, such that, for every domain D and subsets $A, B \subseteq D$, it is true in D iff 'All As are Bs', where the interpretation of P, Q is A, B. It is a classical results (Peters and Wasterstahl [22]) that many generalized quantifiers (e.g. *More than half of*, *Most of* etc.) are not first-order definable. On the other hand, *More than half of* is first-order expressible, in the following sense: for every domain D and subsets $A, B \subseteq D$, there is a first-order sentence $\exists x_{>n}(Px \wedge Qx)$ which is true in D iff 'More than half of As are Bs'. Note that the choice of n must be such that $n = 1/2$ the size of P. Iacona points out that usually it is assumed that, in order to first-order formalize a quantified sentence, the quantifier must be first-order definable. In this sense, 'More than half of the students will fail the class' is not first-order formalizable. However, in a world where the students in question are 20, the sentence can be first-order expressed as 'At least 11 students will fail the class'. Iacona submits that, if we accept that first-order formalization is relative to an interpretation, we should identify formalization with first-order expressibility, and therefore conclude that 'More than half of the students students will fail the class' is formalizable. The same applies to other generalized quantifiers which are not first-order definable, but are first-order expressible.

The focus of Mario Gómez-Torrente's essay, 'Quantifiers and Referential Use', is the attributive vs. referential use distinction in the case of definite descriptions, as well as non-descriptive quantifier phrases. According to the received view, only descriptions can be used referentially in standard settings—for instance, when we utter 'The murderer of Smith is insane' while aiming to refer to a particular person sitting in front of the judge in a court of law. *Contra* the received view, Gómez-Torrente exhibits examples of quantifier phrases being used referentially—for instance, when we utter 'All/most murderers of Smith are insane' and aim to refer to a specific group of people sitting in front of the judge. If Gómez-Torrente is right, we should reassess our explanation of the cases of referential use. Indeed, it has been argued that the distinction between referential and attributive use should be explained semantically, namely as a consequence of an ambiguity in the semantic contribution of ‘the’ (which could act either as an indexical, in the referential case, or as a constituent of a description, in the attributive case). But if Gómez-Torrente’s examples are any evidence, the standard explanation will lose its force, since quantifier phrases do not contain any occurrences of ‘the’.

Gómez-Torrente argues for a pragmatic explanation of the referential use of descriptive and quantifier expressions, in a disjunctive fashion. When the referential use is in a standard setting (‘The/all/most murderers of Smith is/are insane’), the reference to a particular individual, or individuals, is obtained by an inference which obeys the Gricean maxims. When the referential use is in a non-standard setting (‘Most people in this room have played cricket’; ‘The students in my seminar showed up at the party’), the particular reference is obtained by an inference which violates the Gricean maxims.

The third main thread in this volume has to do with *Carnap’s and Quine’s legacy* on quantification theory. In his essay ‘Carnap, Quine, Quantification and Ontology’, Gregory Lavers analyzes the Carnap-Quine debate on quantification and ontological commitment, with a special focus on the mutual misunderstandings. Carnap’s view on ontology shifted quite dramatically from *The Logical Syntax of Language* to *Empiricism, Semantics and Ontology* (Carnap [5, 6]). In the former, he proposed that scientific languages should be studied without employing any ontological vocabulary, such as ‘reference’ and ‘existence’. Accordingly, the statement *‘five’ refers to a number* should be paraphrased as: *‘five’ is a number term*. However, the development of Tarskian semantics had him convinced that we can speak of reference and existence without entering the fray of metaphysical debates. For example, *‘five’ refers to a number* can be demonstrated from:

1. ‘Five’ refers to five
2. Five is a number

where (1) is a Tarskian semantic condition, whereas (2) is provable in a Frege-style background arithmetic theory. As it turns out, argues Lavers, Quine was never able to appreciate Carnap’s transition from *Logical Syntax of Language* to *Empiricism, Semantics and Ontology*.

Neither was Carnap able to fully distinguish Quine’s view from his own. Quine’s view on ontology boils down to the well-known dictum: to be is a value of a bound variable (cf. Parsons [21].). In a less concise form: the entities we should accept are those quantified over by our best theories (where ‘best’ is defined by a number of theoretical virtues). Quine thought that he had found, with that criterion, a way to capture the problem of ontology in a language-independent fashion. Carnap, on the other hand, couldn’t see a substantive difference between Quine’s view and his own later position. But notice that Carnap’s proof that ‘five’ refers to a number depends on the Tarskian condition (1), in which the notion of reference is language-dependent. Consequently, the gap between the two views is wider than Carnap might have thought. In particular, Carnap’s view, unlike Quine’s, entails *quantifier variance*, the view that the meaning of quantifiers is language-dependent, and that the different quantifier meanings have equal metaphysical merit—which is to say, ontological questions can be answered in non-equivalent ways, depending on the language in which they are cast (cf. Chalmers, Manley and Wasserman [8]).

Quantifier variance is the focus of Davis Liebesman's 'Quantifier Variance, Intensionality, and Metaphysical Merit'. The contemporary champion of quantifier variance is Eli Hirsch (whose views are articulated in [14, 15]), who interprets metaphysical merit in terms of expressive power: two languages (with their respective quantifiers) are of equal metaphysical merit if they can express the same contents, where content is construed intensionally. Liebesman offers two arguments against intensional (coarse-grained) content in the definition of metaphysical merit:

1. Quantifier variance is a thesis about unrestricted quantification. But the distinction between restricted and unrestricted quantification is hyperintensional. For example, the quantifier of the mereological nihilist and the universalist quantifier restricted to mereological atoms are intensionally equivalent, and yet only the former is unrestricted. So, in order to properly formulate quantifier variance as the thesis that different *unrestricted* quantifier meanings have equal metaphysical merit, we need a hyperintensional view of content.
2. According to the best known intensional theory of content, which is due to Stalnaker [29, 30], a dispute about whether there are tables is a metasemantic dispute about which proposition is expressed by 'there are tables'—for instance, it could express the proposition that there are tables, or it could express the proposition that there are simples arranged table-wise. However, metasemantic facts are facts about language use, which have no metaphysical merit. So, it is trivial that they have equal metaphysical merit. So, deflationism about existence of tables follows automatically from Stalnaker's theory. However, if that is the case with 'there are tables', it must also be the case with 'there are numbers', which is arguably *not* a verbal dispute. So, Stalnaker's intensional theory of content overgenerates cases of quantifier variance.

Liebesman concludes that Hirsch should adopt a notion of metaphysical merit based on hyperintensional content, and sketches a candidate theory that could make better sense of the notion of quantifier variance.

In 'Chalmers, Quantifier Variance and Mathematicians' Freedom', Sharon Berry argues that quantifier variance could help us understand a peculiar feature of mathematical practice. Berry starts from the observation that we are often struck by mathematicians' freedom to introduce new kinds of entities such as irrational and complex numbers, sets, or specific classes of functions. She argues that this practice is a form of quantifier variance which allows the mathematicians to switch from a language that doesn't quantify over Fs to one in which the quantifier acquires a new meaning that allows entities of kind F in its domain. In order to model quantifier variance, she draws on the proposal of Chalmers [7]. According to him, the quantifier variantist associates to each quantifier meaning a furnishing function mapping worlds to domains. Accordingly, there could be a function, associated to the quantifier of the mereological nihilist, that maps worlds to domains of atoms, and one function which maps worlds to the same respective domains closed under arbitrary fusions. Now, one problem with Chalmers' interpretation of quantifier variance is that the models are set theoretic in nature and so are inadequate for theories that require class-sized models, such as set theory in its standard interpretation. Berry, instead of appealing to a world-based model theory, proposes a way to redefine the quantifiers by means of a primitive notion of logical possibility. For instance, suppose we start with a language $\mathcal{L}_1$ that doesn't allow quantification over numbers. We could then define a more abundant notion of existence in a new language $\mathcal{L}_2$ by means of the following clause: 'there are numbers' is true in $\mathcal{L}_2$ iff 'Necessarily, if Peano Arithmetic is true, then there are numbers' is true in $\mathcal{L}_1$. Since modality is here construed as logical, and therefore broader than metaphysical, even a $\mathcal{L}_1$-speaking nominalist could accept that 'Necessarily, if PA is true, then there are numbers' is non-vacuously true in $\mathcal{L}_1$.

Quine's dictum that to be is to be the value of a bound variable is the target of Aldo Antonelli's 'Life on the Range: Quine's Thesis and Semantic Indeterminacy'. First of all, Antonelli argues that being the value of a bound variable is not a necessary condition for something to have being. In order to show this, he appeals to non-standard (Henkin) models. If quantifiers are interpreted in the usual way as second-order properties, the extension of $\exists$ will be the set of non-empty sets of the model. Now, since the model is non-standard, there will be some plurality of objects, call them the Fs, such that the set of them is not in the model, and therefore $\exists xFx$ will be false. Nevertheless, the Fs will intuitively exist, insofar as they are members of the domain of the model. Antonelli also argues that being the value of a bound variable is not a sufficient condition for something to have being, either. He shows this by assuming Arthur Prior's thesis that if A entails B, then the ontological import of B cannot exceed the ontological import of A (Prior [23]). As a consequence, since Fa entails $\exists xFx$, but Fa doesn't carry ontological commitment to Fs, neither does $\exists xFx$.

To the question 'what is there?', Quine famously replied: 'everything!' (Quine [24]). In his contribution, '"There Is an 'Is' in There Is": Meinongian Quantification and Existence', Francesco Berto contrasts Quine's thesis to Meinong's that there are things that don't exist. In particular, Berto considers two objections that have been raised against Meinong:

Objection from equivocation: The Meinongians are misunderstanding the meaning of some term, presumably the quantifier; therefore, they are unconsciously changing the subject.

Objection from analyticity: Quine's view is analytically true, therefore Meinong's is an analytic falsehood.

Berto replies to the objection from equivocation by pointing out that the thought behind it is that Quine's thesis is epistemically analytic, in such a way that who denies it misunderstands it. (A sentence is epistemically analytic iff understanding it is sufficient for assenting to it. See Williamson [32].) Berto's rejoinder is that there is no evidence that Meinongians are not competent speakers of the relevant fragment of language, containing quantifiers and the existence predicate. In fact, Meninongians had been using such expressions long before becoming Meinongian, and it is implausible to believe that they had suddenly became incompetent at using the relevant terms. The proponent of the objection from equivocation might reply that Quineans and Meinongians are equally competent speakers outside the ontology room, whereas they have a verbal disagreement inside the room. As Berto points out, however, this reply is a special case of a broader strategy which attempts to deflate debates on the nature of logic (e.g., on negation). But usually both parties in each debate agree that they are talking about the same subject matter, and that the disagreement is substantive.

The objection from analyticity depends on two allegedly analytic premises: that quantification ascribes being; and that being is tantamount to existence. Meinongians might resist either, and Berto argues that they should resist the former. Indeed, the Meinongian should point out that quantifier expressions in some languages other than English don't appear to ascribe being in any obvious way (*es gibt* in German, *il'y a* in French), thus there is no evidence that quantifiers in general attribute being.

Quine famously warned us against the practice of quantifying into intensional contexts (Quine [25]). In his essay 'Qualifying Quantifying-in', Bjørn Jespersen considers two Quinean arguments against quantifying-in. One is the *argument from double-think*, which derives a contradiction from some apparently safe premises:

1. Martha believes that the man with the hat is the murderer
2. Martha doesn't believe that her neighbor is the murderer
3. The man with the hat = Martha's neighbor
4. There is someone such that Martha does and doesn't think that he is the murderer

The second Quinean attack on quantifying-in is the *argument from non-factivity*:

1. Tilman is seeking the fountain of youth
2. Therefore, there is something which Tilman is seeking

where a simple application of existential generalization leads from truth to falsity. The standard strategy for blocking the two arguments consists in regimenting intensional contexts by means of some intensional logic. Jespersen's goal is instead to show that Quine's challenge can be accommodated within an extensional logic, as long as we allow quantification over intensions. In order to see how that works, consider the argument from double-think. Surely, in actuality the intensional object *the man with the hat* happens to pick out the same individual as the intensional object *Martha's neighbor*; nevertheless, the two intensional objects are distinct, since they have distinct modal profiles. As a consequence, the two intensional objects, which are the objects of Martha's belief (or lack thereof) cannot be replaced *salva veritate*. Appealing to quantification over intensions can offer a solution to the argument from non-factivity, as well. For what Tilman is seeking is not some non-existent individual, but rather an intension, which exists at all worlds, including those where it doesn't pick anything out. Hence, the correct conclusion of the second argument will be: Therefore, there is some role/intension which Tilman is seeking. In his essay, Jespersen shows how such intuitions can be systematized within an *extensional* logic of intensions.

The fourth main topic of this volume is the role quantification theory has played in reshaping *ontology* and, more generally, *metaphysics*. In his contribution 'Making

Quantified Truths True', Axel Barceló tackles the issue of truthmakers for quantified truths. A truthmaker is that in virtue of which a truthbearer (a proposition, say) is true. For instance, ⟨John is X⟩ is made true by the fact that John is X. It is intuitive to think that truthmaking is a necessary relation, so that in every world where John is X, the truthbearer ⟨John is X⟩ is true. But things aren't so simple. In actuality, ⟨All chimney sweeps are male⟩ is true, and presumably has as its own truthmaker, namely the collection *C* of all actual chimney sweeps. However, in a world containing *C* as well as an extra female chimney sweep, ⟨All chimney sweeps are male⟩ is false. So, it seems that truthmaking is not necessary after all. In order to restore the necessity of truthmaking, some (Beall [2], Armstrong [1]) have suggested that the actual truthmaker of ⟨All chimney sweeps are male⟩ should be *C* plus the fact that there are no other chimney sweeps other than those in C (a totality fact). However, totality facts are negative facts, and the existence of negative facts is highly contentious (Molnar [19]).

Barceló's solution involves construing truthmaking as a contingent relation between truth-maker and truth-bearer. The gist of the proposal is that truthmaking is a defeasible relation: *a* holds *R* defeasibly to *b* if *a* holds *R* to *b* barring any defeating circumstances (defeaters). For example, perceiving some thing *x* as red is defeasible evidence for knowing that *x* is red—in particular, the perception is not evidence for knowing that *x* is red if the subject comes to believe that her perception takes place in an environment lit by red light bulbs (the defeater). The case of quantified truths is analogous. ⟨All chimney sweeps are male⟩ is made true in actuality by *C*, although the truthmaking relation fails in worlds with defeaters, such as one with a female chimney sweep. But the absence of defeaters is not part of the truthmaker of ⟨All chimney sweeps are male⟩, just like the absence of the belief that the perception takes place in an environment lit by red light bulbs is not part of the reasons for knowing that the object *x* is red. It is worth remarking that Barceló's proposal has one major consequence. Truthmaking is often regarded as a quintessential case of grounding (Schaffer [28]): the truthbearer's truth is completely grounded in the truthmaker's existence. But grounding is typically construed as satisfying a necessitation constraint: *x* completely grounds *y* only if the existence of *x* strictly entails the existence of *y* (Fine [10]). But if truthmaking is contingent, then either truthmaking is not an instance of grounding or grounding doesn't satisfy necessitation.

In his essay 'Necessarily Maybe. Quantifiers, Modality and Vagueness', Alessandro Torza works out a model theory for a first-order modal language with determinacy operator. The modal fragment is interpreted via counterpart-theoretic semantics; determinacy and the cognate notion of vagueness are interpreted via supervaluations. So, the framework involves quantification three times over: over world-bound individuals, worlds and precisifications. Torza argues that vagueness of a modal statement could have multiple sources: predication, intensional identity (due to indeterminacy of the counterpart relation) and modality itself (due to indeterminacy of the range of possible worlds). The latter is the most interesting case, since the standard literature on modality implicitly assumes that it is determinate what worlds there are absolutely.[6] In particular, Torza argues that indeterminacy about the range of worlds can arise in modal realism due to cardinality issues. Indeed, David Lewis' *Restricted Modal Recombination* principle entails: for any objects in any worlds, there exists a world that contains any number of duplicates of all those objects, shape and size permitting (Lewis [16]). The 'shape and size' parameter is supposedly determined by some natural break in the mathematical generalization of ordinary spacetime manifolds. Lewis, however, contemplates the possibility that there could be more than one such natural break. In that scenario, there would be multiple, equally natural candidate generalizations of ordinary spacetime manifolds. It would then be indeterminate where *the* break is, although it is determinate that there is one. Suppose for instance that the suitable breaks are defined by the cardinalities k_1 and k_2 (representing sizes of manifolds), for $k_1 < k_2$. Then it will be indeterminate whether there are worlds of size k_2. Let's make the auxiliary assumption that it is determinate what there is at any given world (i.e., no world-bound quantifier vagueness), and therefore that the size of each possible world is determinate. It follows that it is vague what worlds there are, absolutely. For on the precisification of modal space where the break is set by k_2, there will be some world w with k_2 coexistent objects, and since the size of each world is determinate, w will not exist in the precisification of modal space that allows at most k_1 coexistent objects. This conclusion should come as a surprise since Lewis was vocal in denying that absolute quantification could ever be vague (Lewis [16, pp. 212–13]).

The possibility of absolute quantification is the focus of 'Absolute Generality and Semantic Pessimism' by J. P. Studd. Although most instances of quantification are restricted ("I have everything packed"), sometimes we want to quantify over absolutely everything, in particular in the ontology room ("properties/mereological sums do not exist"). It has been argued that absolute generality is incoherent, or at very least hard to capture. For suppose that there is a first-order language quantifying over absolutely everything. Then we could not define its semantics in a standard set-theoretic first-order metalanguage, for the domain would be a set comprising all sets, and Zermelo-Fraenkel set theory entails the non-existence of a universal set. (For an overview of arguments for and against absolute generality, see the introduction to Rayo and Uzquiano [27].) But Williamson [31] has argued that the absolutist has better expressive resources than the relativist. The moral of Williamson's argument is that (i) the absolutist can define truth-conditions for an interpreted language in an absolutist meta-language (roughly: $\forall x Px$ is true iff everything makes Px true); whereas (ii) the relativist cannot define truth-conditions for an interpreted language in a relativist meta-language.

Drawing on Fine [9], Studd distinguishes between *restrictivist relativism* (the universe of discourse of a language can be absolute, but no particular quantifier in that language can encompass the universe) and *expansionist relativism* (a quantifier in a given language can encompass the universe of discourse of that same language,

---

[6] But see Studd's contribution in this volume for a discussion of the problems facing absolute generality.

but the latter cannot encompass the universe of discourse of every other language). Studd argues that Williamson's argument for (ii) is unsound for the case of expansionist relativism, and therefore that we don't have a conclusive semantic case in favor of absolutism.

Alberto Voltolini's 'What's in a (mental) picture' submits an original solution to the classical problem of characterizing the relation of *inexistence* holding between an intentional object and a mental state. One fairly standard proposal is to construe inexistence in terms of narrow scope existence: intentional object a exists in mental state S iff, according to S, a exists. Notice that this view does not entail any ontological commitment relative to the object a, since 'a exists' is in the scope of the according-to-S operator. This feature is an obvious virtue when a is a non-existent or fictional object. However, Voltolini objects to the view by pointing out that in some cases the quantifier in 'a exists' should take wide scope, namely in those cases where we have a thought about an existing object. In such cases, the narrow scope existence view appears to be inadequate.

Another proposal hinges on the mind dependence of intentional objects. On this view, intentional object a exists in mental state S iff necessarily, if a exists then S exists. But, as Voltolini correctly points out, this view leads to some sort of 'voodoo metaphysics' in which my mother-in-law would disappear if I had never thought of her.

Voltolini's own proposal is to turn the tables on the mind dependence view. He submits the following characterization of inexistence in terms of constitution: a exists in mental state S iff (1) a individuates S and (2) necessarily, a is part of S, if S exists. What follows is in fact a Meinongian theory of intentional objects. For if the existence of my act of thinking of Sherlock Holmes is identified by Sherlock Holmes, then the latter's existence is presupposed by the existence of my own mental act. Interestingly enough, Voltolini's theory appears therefore to converge with Moltmann's linguistic argument in favor of non-existent objects.

The fifth and last part of this volume deals with issues in quantification theory within specific *logical systems*. In his essay 'What's So Bad about Second-Order Logic?', Jason Turner tackles the fairly widespread view that second order logic (SOL) is not *logic*. One influential line of thought, reminiscent of Quine's complaint that SOL is set theory in sheep's clothing, is that SOL is not ontologically innocent (cf. Quine [26]). For second-order quantifiers must be interpreted by reference to property-like objects, and logic is not supposed to involve commitment to such entities. One possible way to avoid ontological commitment to property-like entities is to interpret second-order quantifiers adverbially, instead of objectually. Accordingly, '$\exists X \ldots$' would be interpreted as 'there are things related somehow that…'. This strategy, however, seems applicable only in a restricted class of cases. A different strategy is to implicitly characterize second-order quantifiers via their theoretical role, rather than by their reference to any objects. One limitation of this approach is that SOL is not axiomatizable, and therefore no complete characterization of SOL can be provided.

A further issue is that there exists a formula α of SOL which is valid iff the Continuum Hypothesis (CH) is true. This fact is *prima facie* incompatible with two features of logic. The first feature is the topic neutrality of logic. One could rejoin, however, that there are also validities in FOL which depend on the features of first-order models, and yet we usually take FOL to be logic. So, it is not clear that the kind of dependence between validities and model-theoretic or set-theoretic truths is in tension with the topic neutrality of logic. The second feature is a normativity condition: it is not rational to reject logical validities—and yet it seems rational to reject the truth of CH. Turner, however, points out that the claim that SOL is logic can be interpreted as the claim that SOL is sound and complete with respect to some intuitive notion of logical consequence ($\vdash_L$), which is to say: $\Delta \vdash_{SOL} \phi \Rightarrow \Delta \vdash_L \phi \Rightarrow \Delta \models_{SOL} \phi$. But we know that proof-theoretic SOL consequence ($\vdash_{SOL}$) is strictly stronger than model-theoretic SOL consequence ($\models_{SOL}$), due to incompleteness. So, intuitive logical consequence ($\vdash_L$) could lie between the two technical notions of second-order consequence. In particular, if $\vdash_L$ is strictly stronger than $\models_{SOL}$, it will follow that the second-order sentence α equivalent to CH is second-order valid although not intuitively valid. So we wouldn't have any normative commitment to the truth of CH.

In his contribution '$\forall$ to ω', Elia Zardini elaborates on his previous work, where he articulated a theory of vagueness to solve the sorites paradox, and a theory of truth to solve the liar's paradox. As it turns out, the classical rule of universal generalization fails in both theories. Zardini sees the root of the failure in a distinction between 'anything' and 'everything', which is blurred in the classical case due to the standard rule of universal generalization which permits to infer $\forall xPx$ from Py, as long as y doesn't occur free in any of its premises. A typical case where the anything vs. everything distinction shows up are free-choice permissions: it is possible for someone to choose any main course in a menu, but not possible to choose every main course in the menu. According to Zardini, a "natural, innovative strategy to account for this contrast is to postulate that the proposition that one has any main course is weaker than the proposition that one has every main course, and so that permission of the former does not entail permission of the latter". After rejecting the everything-from-anything inference, Zardini proposes a new universal generalization rule: $Pt_1 \ldots Pt_n \vdash \forall xPx$, where $t_1, \ldots, t_n$ are all the individual terms in the language. This rule leads to a solution to free-choice permission puzzles. First, notice that: one has main course 1,..., one has main course $n \vdash$ one has every main course. So, in order to get 'possibly, one has every main course', we need 'possibly, one has main course 1,... one has main course n', which is false. For all we have is the weaker 'possibly, one has main course 1,..., possibly, one has main course n'. Zardini considers and rejects a number of objections to his universal generalization rule—for instance, that it is unsound if not everything in the domain is named. But Zardini points out that this would be also an objection to classical universal instantiation, and therefore his non-classical rule doesn't fare worse than the classical case in that respect.

If quantifiers are variable-binding operators, there are operators that are not variable-binding and nevertheless show quantificational behavior, in the sense that, from a model-theoretic point of view, range over a domain of objects. A typical case of such non-overtly quantificational operators are intensional operators—modal (ranging over worlds), epistemic (ranging over informational states), temporal (ranging over times) etc. Tero Tulenheimo focuses on the temporal case in his 'Cross-World Identity, Temporal Quantifiers and the Question of Tensed Contests'. He draws on the work of Jaakko Hintikka who, in the case of epistemic logic, distinguishes between physical vs. perspectival individuation of objects. The difference is captured by means of distinct functions from epistemic worlds to objects (Hintikka [12, 13]). Tulenheimo generalizes Hintikka's idea to the temporal case, defining a model theory for tensed languages featuring one function from worlds to instants which corresponds to objective ('physical') time, and as many such functions as there are agents, the perspectival ('subjective') times. In defining the semantics, Tulenheimo lets the temporal operators range not over instants of time, but rather over functions from worlds to instants. This machinery allows us to represent some ambiguities in the way we can understand the content of a tensed assertion. Take for example 'Mary knows that it is now 4.30', and suppose that the time of utterance is indeed 4.30. The utterance can be interpreted in two ways. First, it could mean that (1) Mary knows the intended use of the expression '4.30'. Model theoretically: in every world compatible with Mary's knowledge, the term '4.30' refers to 4.30, i.e. the instant picked out by the objective time-function *now*. Alternatively, 'Mary knows that it is now 4.30' could have the more standard meaning that (2) Mary knows that the time that she perceives as present is 4.30. Model-theoretically: in every world compatible with Mary's knowledge, the term '4.30' refers to Mary's subjective present, i.e. the instant picked out by the perspectival term-function *Mary's now*.

References

1. Armstrong, D. 2004. *Truth and truthmakers*. Cambridge: Cambridge University Press.
2. Beall, J.C. 2000. On truthmakers for negative truths. *Australasian Journal of Philosophy* 78: 264–268.
3. Beth, E.W. 1953. On Padoa's method in the theory of definition. *Indagationes Mathematicae* 15: 330–339.
4. Beth, E.W. 1964. *The foundations of mathematics*. Amsterdam: North-Holland. 1959, Harper Torch books: 290–293.
5. Carnap, R. 1934/1937. *The logical syntax of language*. London: Routledge & Kegan Paul.
6. Carnap, R. 1991. Empiricism, semantics, and ontology. In *The Philosophy of Science*, ed. R. Boyd, P. Gasper, and J.D. Trout, 85–98. Cambridge: MIT.
7. Chalmers, D. 2009. Ontological anti-realism. In *Metametaphysics: New essays on the foundations of ontology*, ed. R. Wasserman, D. Chalmers, and D. Manley. Oxford: Oxford University Press.
8. Chalmers, D., D. Manley, and R. Wasserman. (eds.) 2009. *Metametaphysics. New essays on the foundations of ontology*. Oxford: Oxford University Press.
9. Fine, K. 2006. Relatively unrestricted quantification. In: *Absolute generality*, ed. A. Rayo and G. Uzquiano, 20–44. Oxford: Clarendon Press.
10. Fine, K. 2012. Guide to ground. In *Metaphysical grounding: Understanding the structure of reality*, ed. F. Correia and B. Schnieder, 37–80. Cambridge: Cambridge University Press.

11. Gómez-Torrente, M. 2002. The problem of logical constants. *The Bulletin of Symbolic Logic* 8(1): 1–37.
12. Hintikka, J. 1969. *Models for modalities*. Dordrecht: Reidel.
13. Hintikka, J. 1975. *The intentions of intentionality and other new models for modalities*. Dordrecht: Reidel.
14. Hirsch, E. 1993. *Dividing reality*. Oxford: Oxford University Press.
15. Hirsch, E. 2011. *Quantifier variance and realism. Essays in metaontology*. New York: Oxford University Press.
16. Lewis, D. 1986. *On the plurality of worlds*. Oxford/New York: Blackwell.
17. Lindström, P. 1966. First order predicate logic with generalized quantifiers. *Theoria* 32: 186–195.
18. McGinn, C. 2000. *Logical properties*. Oxford: Oxford University Press.
19. Molnar, G. 2000. Truth makers for negative truths. *Australasian Journal of philosophy* 78(1): 72–86.
20. Montague, R. 1973. The proper treatment of quantification in ordinary English. In *Approaches to natural language*, ed. J. Hintikka, J. Moravcsik, and P. Suppes, 221–242. Dordrecht: Reidel.
21. Parsons, T. 1970. Various extensional notions of ontological commitment. *Philosophical Studies* 21(5): 65–74.
22. Peters, S., and D. Westerståhl. 2006. *Quantifiers in language and logic*. Oxford: Clarendon Press.
23. Prior, A.N. 1971. *Objects of thought*, ed. P.T. Geach and A.J.P. Kenny, ix–175. Oxford: Clarendon Press.
24. Quine, W.V.O. 1948. On what there is. *The Review of Metaphysics* 2: 21–38.
25. Quine, W.V. 1956. Quantifiers and propositional attitudes. *Journal of Philosophy* 53: 177–187.
26. Quine, W.V.O. 1970. *Philosophy of logic*. Englewood Cliffs: Prentice-Hall.
27. Rayo, A., and G. Uzquiano. 2006. *Absolute generality*. Oxford: Oxford University Press.
28. Schaffer, J. 2008. Truth-maker commitments. *Philosophical Studies* 141: 7–19.
29. Stalnaker, R. 1978. Assertion. In *Syntax and semantics 9*, ed. P. Cole. New York: Academic.
30. Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT.
31. Williamson, T. 2003. Everything. *Philosophical Perspectives* 17(1): 415–465.
32. Williamson, T. 2007. *The philosophy of philosophy*. Oxford: Blackwell.
33. Zucker, J.I. 1978. The adequacy problem for classical logic. *Journal of Philosophical Logic* 7: 517–535.

Part I
Logical Constants

Chapter 2
Which Quantifiers Are Logical? A Combined Semantical and Inferential Criterion

Solomon Feferman

Abstract The aim of logic is to characterize the forms of reasoning that lead invariably from true sentences to true sentences, independently of the subject matter; thus its concerns combine semantical and inferential notions in an essential way. Up to now most proposed characterizations of logicality of sentence generating operations have been given either in semantical or inferential terms. This paper offers a combined semantical and inferential criterion for logicality (improving one originally proposed by Jeffery Zucker) and shows that any quantifier that is to be counted as logical according to that criterion is definable in first order logic.

The aim of logic is to characterize the forms of reasoning that lead invariably from true sentences to true sentences, independently of the subject matter. The sentences involved are analyzed according to their logical (as opposed to grammatical) structure, i.e. how they are compounded from their parts by means of certain operations on propositions and predicates, of which the familiar ones are the connectives and quantifiers of first order logic. To spell this out in general, one must explain how the truth of compounds under given operations is determined by the truth of the parts, and characterize those forms of rules of inference for the given operations that insure preservation of truth. The so-called problem of "logical constants" (Gómez-Torrente [8]) is to determine all such operations. That has been pursued mostly via purely semantical (qua set-theoretical) criteria on the one hand–stemming from Tarski [17]–and purely inferential criteria on the other–stemming from Gentzen [6] and pursued by Prawitz [14], among others–even though on the face of it a combination of the two is required.[1] What is offered here

[1]Some further contributions to the semantical approach are Sher [16] and McGee [12], and to the inferential approach is Hacking [9]; Gomez-Torrente [8] provides a useful survey of both approaches. I have critiqued the semantical approach as given by set-theoretical criteria in Feferman [4, 5] where, in conclusion, I called for some combined criterion.

S. Feferman (✉)
Department of Mathematics, Stanford University, Stanford, CA 94305-2125, USA
e-mail: feferman@stanford.edu

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_2

is such a combined criterion for quantifiers, whose semantical part is provided by Lindström's [11] generalization of quantifiers, and whose inferential part is closely related to one proposed by Zucker [18].[2] On the basis of this criterion it is shown that any quantifier that is to be counted as logical is definable in classical first order logic (FOL). In addition, part of the proof idea is the same as that provided by Zucker, but his proof itself needs to be corrected in at least one essential respect that will be explained below; fixing that up is my main contribution here in addition to elaborating the criterion for logicality. One basic conceptual difference that I have with Zucker is that he regards the meaning of a quantifier to be given by some axioms and rules of inference, provided those uniquely determine it on an inferential basis, whereas I assume that its meaning is specified semantically; that is the viewpoint both of workers in model-theoretic logics (cf. Barwise and Feferman [1]) and of workers on quantifiers in natural language (cf. Peters and Westerståhl [13]). For Zucker's point of view, see the Discussion below.

Given a non-empty universe of discourse U and $k \geq 1$, a k-ary relation on U is simply a subset P of U^k; we may also identify such with k-ary "propositional" functions $P : U^k \to \{t, f\}$, where t and f are the truth values for truth and falsity, respectively. $P(x_1, \ldots, x_k)$ may thus be read as "P holds of $(x_1, \ldots, x_k)$" or as "$P(x_1, \ldots, x_k)$ is true." Q is called a *(global) quantifier* of type $\langle k_1, \ldots, k_n \rangle$ if Q is a class of relational structures of signature $\langle k_1, \ldots, k_n \rangle$ closed under isomorphism. A typical member of Q is of the form $\langle U, P_1, \ldots, P_n \rangle$ where U is non-empty and P_i is a $k_i - ary$ relation on U. Given Q, with each U is associated the *(local) quantifier* Q_U on U which is the relation $Q_U(P_1, \ldots, P_n)$ that holds between $P_1, \ldots, P_n$ just in case $\langle U, P_1, \ldots, P_n \rangle$ is in Q. Alternatively we may identify Q_U with the associated functional from propositional functions of the given arities on U to $\{t, f\}$. Examples of such quantifiers may be given in set-theoretical terms without restriction. Common examples are the uncountability quantifier of type $\langle 1 \rangle$, the equi-cardinality quantifier of type $\langle 1, 1 \rangle$, and the "most" quantifier of type $\langle 1, 1 \rangle$. However, even though the definitions of those refer to the totality of relations of a certain sort (namely 1-1 functions), all quantifiers in Lindström's sense satisfy the following principle:

Locality Principle. Whether or not $Q_U(P_1, \ldots, P_n)$ is true depends only on U and $P_1, \ldots, P_n$, and not on what sets and relations exist in general over U.

As shown by Lindström, given any first-order language L with some specified vocabulary of relations, functions and constant symbols, we may add Q as a formal symbol Q to be used as a new constructor of formulas ϕ from given formulas ψ_i, $i = 1, \ldots, n$. For each i, let $\underline{x}_i$ be a $k_i - ary$ sequence of distinct variables such that

[2] An unjustly neglected paper, along with Zucker and Tragesser [19].

$\underline{x}_i$ and $\underline{x}_j$ are disjoint when $i \neq j$, and let $\underline{y}$ be a sequence of distinct variables disjoint from all the $\underline{x}_i$. The syntactical construction associated with Q takes the form

$$\phi(\underline{y}) = \mathrm{Q}\underline{x}_1 \ldots \underline{x}_n(\psi_1(\underline{x}_1, \underline{y}), \ldots, \psi_n(\underline{x}_n, \underline{y}))$$

where the $\underline{x}_i$ are all bound and the free variables of ϕ are just those in $\underline{y}$. The satisfaction relation for such in a given L-model $\mathcal{M}$ is defined recursively: for an assignment $\underline{b}$ to $\underline{y}$ in U, $\phi(\underline{b})$ is true in $\mathcal{M}$ iff $(U, P_1, \ldots, P_n)$ is in Q when each P_i is taken to be the set of k_i-tuples $\underline{a}_i$ satisfying $\psi_i(\underline{a}_i, \underline{b})$ in $\mathcal{M}$.

Next what is needed to bring inferential considerations into play is to explain which quantifiers have axioms and rules of inference that completely govern its forms of reasoning. It is here that we connect up with the inferential viewpoint, beginning with Gentzen [6]. Remarkably, he showed how *prima facie* complete inferential forms could be provided separately for each of the first-order connectives and quantifiers, whether thought of constructively or classically, via the *Introduction and Elimination Rules* in the calculi NJ and NK, resp., of natural deduction. In addition, he first formulated the idea that the *meaning* of each of these operations is given by their characteristic inferences. Actually, Gentzen claimed more: he wrote that "the [Introduction rules] represent, as it were, the 'definitions' of the symbols concerned." (Gentzen [6], p. 80). Prawitz put teeth into this by means of his Inversion Principle (Prawitz [14], p. 33): namely, it follows from his normalization theorems for NJ and NK that each Elimination rule for a given operation in either calculus can be recovered from the appropriate one of its Introduction rules when that is the last step in a normal derivation.

As I have stated above, in my view the meaning of given connectives and quantifiers is to be established semantically in one way or another *prior* to their inferential role. Their meanings may be the primitives of our reasoning in general, including "and", "or", "not", "if ... then", "all", "some"–or they may be understood informally like "most", "has the same number as", etc., in a way that may be explained precisely in basic mathematical terms. What is taken from the inferentialists (or Zucker) is not the thesis as to meaning but rather their formal analysis of the essential principles and rules which are in accord with the prior semantical explanations and that govern their use in reasoning. And in that respect, *the Introduction and Elimination Rules for each logical operation of first-order logic implicitly characterize it in the sense that any other operation satisfying the same rules is provably equivalent to it.*[3] That unicity will be a key part of our criterion for logicality in general.

[3]The observation that the natural deduction Introduction and Elimination rules for the operations of FOL serve to uniquely specify each such operation is, I think, well known. At any rate, one can find it stated in Zucker and Tragesser [19] p. 509. In apparent agreement with Gentzen that the Introduction rules provide the meaning of each operation, they say that the related Elimination rules serve to "stabilize" or "delimit" it.

To illustrate, since I will be dealing here only with classical truth functional semantics, I consider schematic axioms and rules of inference for sequents $\Gamma \vdash \Delta$ as in LK, but in the case of each connective or quantifier, show only those formulas in Γ and Δ directly needed to characterize the operation in question. That may include possible additional side formulas (or parameters), to which all further formulas can be adjoined by thinning. In LK, the *Right* and *Left Introduction Rules* take the place of the *Introduction* and *Elimination Rules*, resp., in NK. I shall then show how unicity is expressed for the corresponding Hilbert-style axioms and rules.

Consider for illustrative purposes the (axioms and) rules for $\to$ and $\forall$. For notational simplicity, $\Rightarrow$ is used for inference from one or more sequents as hypotheses, to a sequent as conclusion.

$$(\mathrm{R}\to)\ r, p \vdash q \Rightarrow r \vdash p \to q \qquad (\mathrm{L}\to)\ p, p \to q \vdash q$$
$$(\mathrm{R}\forall)\ r \vdash p(a) \Rightarrow r \vdash \forall x p(x) \qquad (\mathrm{L}\forall)\ \forall x p(x) \vdash p(a).$$

Given an operation $\to'$ satisfying the same rules as for $\to$ we can infer from the left rule $p \to q, p \vdash q$ the conclusion $p \to q \vdash p \to' q$ by the substitution of $p \to q$ for r in $(\mathrm{R} \to')$; the reverse holds by symmetry. In the case of the universal quantifier, given $\forall'$ that satisfies the same rules as $\forall$, we can derive $\forall x p(x) \vdash \forall' x p(x)$ by substituting $\forall x p(x)$ for r in $(\mathrm{R}\forall')$. What is crucial in these proofs of uniqueness is the use of substitution of the principal formula ($p \to q$ and $\forall x p(x)$ and their $'$ versions, resp.) for a side formula (parameter) r.

If we accept $\to$ as a basic fully understood operator, we can pass to the Hilbert-style axioms and rules for the universal quantifier by simply replacing the turnstile symbol by '$\to$', as follows:

$$(\mathrm{R}\forall)^{\mathrm{H}}\ r \to p(a) \Rightarrow r \to \forall x p(x) \qquad (\mathrm{L}\forall)^{\mathrm{H}}\ \forall x p(x) \to p(a).$$

Then in a suitable metatheory for axioms and rules in which we take *all* the connectives and quantifiers of FOL for granted, we can represent this rule and axiom by the following single statement in which we treat universal quantification as a quantifier Q of type $\langle 1 \rangle$:

$$A(\mathbf{Q})\ \forall p \forall r \{[\forall a(r \to p(a)) \to (r \to \mathbf{Q}(p))] \wedge [\forall a(\mathbf{Q}(p) \to p(a))]\},$$

where 'r' ranges over arbitrary propositions and 'p' over arbitrary unary predicates. Then, as above, we easily show that

$$(A(\mathbf{Q}) \wedge A(\mathbf{Q}')) \to (\mathbf{Q}(p) \leftrightarrow \mathbf{Q}'(p)).$$

Our question now is: Which quantifiers Q in general have formal axioms and rules of inference that uniquely characterize it in the same way as for universal quantification? The answer to that will initially be treated via a *second-order language* L_2 of individuals, propositions and predicates, first without and then with a symbol for Q.

L_2 is specified as follows:

Individual variables: $a, b, c, \ldots, x, y, z$
Propositional variables: $p, q, r, \ldots$ Predicate variables, k-ary: $p^{(k)}$, $q^{(k)}$, $\ldots$; the superscript k may be dropped when determined by context.
Propositional terms: the propositional variables p, q, $r, \ldots$ and the $p^{(k)}(x_1, \ldots, x_k)$ (any sequence of individual variables)
Atomic formulas: all propositional terms
Formulas: closed under $\neg$, $\wedge$, $\rightarrow$, $\forall$ applied to individual, propositional and predicate variables. (Other connectives and quantifiers defined as usual.)

Next, models $\mathcal{M}_2$ of L_2 are specified as follows:

(i) Individual variables range over a non-empty universe U
(ii) Propositional variables range over $\{t,f\}$ where $t \neq f$.
(iii) Predicate variables of k arguments range over $Pred^{(k)}(\mathcal{M}_2)$, a subset of $U^k \rightarrow \{t,f\}$.

Clause (iii) is in accord with the Locality Principle, according to which predicate variables may be taken to range over any subset of the totality of k-ary relations on U.

Satisfaction of a formula ϕ of L_2 in $\mathcal{M}_2$ at an assignment σ to all variables, $\mathcal{M}_2 \vDash \phi[\sigma]$, is defined inductively as follows:

1. For $\phi \equiv p$, a propositional variable, $\mathcal{M}_2 \vDash \phi[\sigma]$ iff $\sigma(p) = t$
2. For $\phi \equiv p(x_1, \ldots, x_k), p$ a k-ary predicate variable, $\mathcal{M}_2 \vDash \phi[\sigma]$ iff $\sigma(p)(\sigma(x_1), \ldots, \sigma(x_k)) = t$
3. Satisfaction is defined inductively as usual for formulas built up by $\neg$, $\wedge$, $\rightarrow$, $\forall$, given the specified ranges in (ii) and (iii) for the propositional and predicate variables when it comes to quantification.

Now, given a quantifier Q of arity $\langle k_1, \ldots, k_n \rangle$, the language $L_2(\mathbf{Q})$ adjoins a corresponding symbol $\mathbf{Q}$ to L_2. This is used to form propositional terms $\mathbf{Q}(p_1, \ldots, p_n)$ where p_i is a k_i-ary variable. Each such term is then also counted as an atomic formula of $L_2(\mathbf{Q})$, with formulas in general generated as before. A model $(\mathcal{M}_2, Q|\mathcal{M}_2)$ of $L_2(\mathbf{Q})$ adjoins a function $Q|\mathcal{M}_2$ as the interpretation of $\mathbf{Q}$, with $Q|\mathcal{M}_2 : Pred^{(k_1)}(\mathcal{M}_2) \times \ldots \times Pred^{(k_n)}(\mathcal{M}_2) \rightarrow \{t,f\}$.

Axioms and rules for a quantifier Q as in LK can now be formulated directly by a sentence A($\mathbf{Q}$) in the language $L_2(\mathbf{Q})$, as was done above for the universal quantifier, by using the associated Hilbert-style rules as an intermediate auxiliary. To formulate the translation in general if we start with rules in the sequent calculus, suppose those for a formal quantifier $\mathbf{Q}(p_1, \ldots, p_n)$ of the sort we are considering are Rule$_1$,..., Rule$_m$, where each Rule$_j$ has 0 or more sequents $\Gamma_{j,v} \vdash \Delta_{j,v}$ in the hypothesis and one sequent $\Gamma_j \vdash \Delta_j$ as conclusion. Some of these will be Right rules and some Left rules for $\mathbf{Q}$.[4] Consider any such Rule$_j$. If there is more

[4] Zucker and Tragesser [19], pp. 502–03 make further assumptions about the nature of the rules in a natural deduction calculus for a candidate operator. Since our criterion will be formulated under much looser assumptions, we don't have to invoke those here.

than one term in the antecedent of one of the sequents in the hypothesis, replace that by their conjunction, and if in the succeedent by their disjunction. Replace an empty antecedent by $\forall p(p \to p)$ and an empty succeedent by $\neg\forall p(p \to p)$. Finally, replace $\vdash$ by $\to$. Next, for each j, take the conjunction of the translations of the $\Gamma_{j,v} \vdash \Delta_{j,v}$, and universally quantify that by all the individual variables that occur in it; call that H_j. Similarly, replace the conclusion $\Gamma_j \vdash \Delta_j$ by the universal quantification C_j over the individual variables of its translation. Finally, replace the inference sign $\Rightarrow$ from the hypotheses to the conclusion by $\to$. Let $B_j = H_j \to C_j$ be the translation of Rule$_j$ thus obtained. Finally, take A(**Q**) to be the sentence

$$\forall p \forall q \forall r \ldots (B_1 \wedge \ldots \wedge B_m),$$

where p, q, r, ... are all the propositional and predicate variables that appear in any of the B_j. Now the criterion for accepting a quantifier Q given by such rules is that they implicitly define Q_U in each model of A(**Q**) (more precisely, the restriction of Q_U to the predicates of the model).

We need not restrict to such specific descriptions of axioms and rules of inference for a global quantifier Q in formulating the following more general partial criterion for acceptance of Q as logical. The reason this is not claimed to be a necessary and sufficient condition for logicality will be discussed below.

Semantical-Inferential Necessary Criterion for Logicality. A global quantifier Q of type $\langle k_1, \ldots k_n \rangle$ is logical only if there is a sentence A(**Q**) in L$_2$(**Q**) such that for each model $\mathcal{M}_2 = (U, \ldots)$, Q_U is the unique solution of A(**Q**) when restricted to the predicates of $\mathcal{M}_2$.

Remark. I spoke above of the use of axioms and rules of inference for a quantifier Q that completely govern its forms of reasoning. One should be careful to distinguish completeness of a system of axioms in the usual sense from completeness of a sentence A(**Q**) for Q in the sense that it meets the above criterion. For example, let Q_α be the type $\langle 1 \rangle$ quantifier which holds of a subset P of U just in case P is of cardinality at least $\aleph_\alpha$. Keisler [10] has proved completeness of a system of axioms for first-order logic extended by Q_1. But it is easily seen that those same axioms are satisfied by Q_α for any α greater than 1 (cf. ibid, p. 29). Hence a sentence A(**Q**) formally expressing Keisler's axioms does not meet the above criterion.

Main Theorem. Suppose Q is a quantifier that satisfies the preceding partial criterion for logicality. Then Q is equivalent to a quantifier defined in FOL.

The sketched proof of the related theorem in Zucker [18] pp. 526ff makes use of a different second order language than here, and claims to apply Beth's definability theorem to obtain an equivalence of **Q** with a formula in FOL. The first problem with that is the question of the applicability of Beth's theorem to a second-order language. That may be possible for certain languages such as L$_2$ whose semantics is not the standard one but rather is "Henkin" or "general". So far as I know a Beth theorem for such has not been established in the literature, even though that is quite plausible. In order to do that, one might try to see how the extant model-theoretic or proof-theoretic proofs can be adapted to such languages. But even if one has done that, all that the corresponding Beth theorem would show is that Q is definable by a formula in L_2; in order to obtain a definition in FOL, one would still have to eliminate the propositional and predicate variables, and that requires a further argument, not considered at all by Zucker. It is shown here how to take care of both difficulties by simulating the languages L_2 and $L_2(\mathbf{Q})$ and their models in corresponding *first-order* languages L_1 and $L_1(\mathbf{Q})$ in which the proposition and predicate variables are taken to be two new sorts of variables at type level 0 besides the individual variables.

Here is the specification of this first-order language L_1:

Individual variables: $a, b, c, \ldots, x, y, z$
Propositional variables: $p, q, r, \ldots$
Propositional constants: $\boldsymbol{t}, \boldsymbol{f}$
Predicate variables $p^{(k)}$ of k arguments for $k \geq 1$; where there is no ambiguity, we will drop the superscripts on these variables.
Predicate constants $\mathbf{t}^{(k)}$ of k arguments for each $k \geq 1$.
In addition, L_1 has for each k a $k+1$-ary function symbol App_k for application of a k-ary predicate variable $p^{(k)}$ to a k-termed sequence of individual variables $x_1, \ldots, x_k$; we write $p^{(k)}(x_1, \ldots, x_k)$ for $App_k(p^{(k)}, x_1, \ldots, x_k)$.
The terms of L_1 are the variables and constants of each sort, as well as the terms $p^{(k)}(x_1, \ldots, x_k)$ of propositional sort for each k-ary predicate variable $p^{(k)}$. The atomic formulas are just those of the form $\pi_1 = \pi_2$, where π_1 and π_2 are terms of propositional sort. Formulas in general are built up from these by means of the first-order connectives and quantifiers over each of the sorts of variables as usual.

By the language $L_1(\mathbf{Q})$ is meant the extension of L_1 by a function symbol $\mathbf{Q}$ taking a sequence $(p_1, \ldots, p_n)$ of predicate variables (not necessarily distinct) as arguments, where p_i is k_i-ary, to a term $\mathbf{Q}(p_1, \ldots, p_n)$ of propositional sort. For any term π of propositional sort, whether in the base language or this extension, we write $T(p)$ for $p = t$, to express that p is true.

The following is a base set S of axioms for L_1:

(i) $\sim(\boldsymbol{t} = \boldsymbol{f})$
(ii) $\forall p(p = \boldsymbol{t} \vee p = \boldsymbol{f})$, ('$p$' a propositional variable)
(iii) $\forall x_1 \ldots \forall x_k(\boldsymbol{t}^{(k)}(x_1, \ldots, x_k) = \boldsymbol{t})$ for each $k \geq 1$
(iv) $\forall p \forall q[\forall x_1 \ldots \forall x_k(p(x_1, \ldots, x_k) = q(x_1, \ldots, x_k)) \rightarrow p = q]$.

The last of these is of course just Extensionality for predicates.

Models $\mathcal{M}_1$ of S are given by any non-empty universe of individuals U as the range of the individual variables, and the set $\{t,f\}$ (with $t \neq f$) as the range of the propositional variables. Furthermore each assignment to a k-ary predicate variable in $\mathcal{M}_1$ determines a propositional function P from U^k to $\{t,f\}$ as its extension, via the interpretation of the application function App_k. By Extensionality, we may think of the interpretation of the k-ary predicate variables in $\mathcal{M}_1$ as ranging over some collection of k-ary propositional functions. The interpretation of $\mathbf{t}^{(k)}$ is just the constant propositional function $\lambda(x_1, \ldots, x_k).t$ on U^k. In the following, all structures $\mathcal{M}_1$ considered are assumed to be models of S.

Each model $\mathcal{M}_2$ of the second order language L_2 may equally well be considered to be a model $\mathcal{M}_1$ of the first order language L_1 in the obvious way. Conversely, by extensionality each of the models $\mathcal{M}_1$ for L_1 may be construed to be a model $\mathcal{M}_2$ for L_2. The essential difference lies in the way that formulas are formed and hence with how satisfaction is defined. In the first-order language, propositional terms are merely such, while they have also been taken to be atomic formulas in the second order language. Recall the abbreviation $T(p)$ for $p = \boldsymbol{t}$ in L_1. Note that any assignment to the variables of L_2 in $\mathcal{M}_2$ counts equally well as an assignment to the variables of L_1 in $\mathcal{M}_1$. All of this goes over to the languages extended by $\mathbf{Q}$ and the corresponding interpretations of it in the respective models.

We define the translation of each formula A of the 2nd order language L_2, with or without $\mathbf{Q}$, into a formula $A\downarrow$ of the 1st order language L_1 by simply replacing each atomic formula τ of A (i.e. each propositional term) by $T(\tau)$. Thus, for example, the translation of the above formula characterizing the axiom and rule for universal quantification is simply

$$\forall p \forall r\{[\forall a(T(r) \to T(p(a))) \to (T(r) \to T(\mathbf{Q}(p)))] \wedge \forall a[T(\mathbf{Q}(p)) \to T(p(a))]\}.$$

Similarly, we obtain an inverse translation from any 1st order formula B of L_1 into a 2nd order formula $B\uparrow$ of L_2 by simply removing each occurrence of 'T' that is applied to propositional terms. The atomic formulas $\pi_1 = \pi_2$ are replaced by $\pi_1 \leftrightarrow \pi_2$. These translations are inverse to each other (up to provable equivalence) and the semantical relationship between the two is given by the following lemma, whose proof is quite simple.

Lemma. *Suppose $\mathcal{M}_1$ and $\mathcal{M}_2$ correspond to each other in the way described above. Then*

(i) If A is a formula of L_2 and σ is an assignment to its free variables in $\mathcal{M}_2$ then $\mathcal{M}_2 \models A[\sigma]$ iff $\mathcal{M}_1 \models A\downarrow[\sigma]$;
(ii) Similarly, if B is a formula of L_1 and σ is an assignment to its free variables in $\mathcal{M}_1$ then $\mathcal{M}_1 \models B[\sigma]$ iff $\mathcal{M}_2 \models B\uparrow[\sigma]$.

Moreover, the same equivalences hold under the adjunction of $\mathbf{Q}$ *throughout.*

Now to prove the main theorem above, suppose $A(\mathbf{Q})$ is a sentence of $L_2(\mathbf{Q})$ such that over each model $\mathcal{M}_2$, Q_U is the unique operation restricted to the predicates of $\mathcal{M}_2$ that satisfies $A(\mathbf{Q})$. Then it is also the unique operation that satisfies $A(\mathbf{Q})\downarrow$ in $\mathcal{M}_1$. So now by the completeness theorem for many-sorted first-order logic, we have provability of

$$(A(\mathbf{Q})\downarrow \wedge A(\mathbf{Q}')\downarrow) \to (\mathbf{Q}(p_1, \ldots, p_n) = \mathbf{Q}'(p_1, \ldots, p_n))$$

in FOL, so that by Beth's definability theorem, which follows from the interpolation theorem for many-sorted logic (Feferman [2]), the relation $\mathbf{Q}(p_1, \ldots, p_n) = \boldsymbol{t}$ is equivalent to a formula $B(p_1, \ldots, p_n)$ of L_1. Moreover, by assumption, in each model $\mathcal{M}_1$, B defines the relation Q_U restricted to the range of its predicate variables (considered as relations). Though B is a formula of L_1, it is not necessarily first-order in the usual sense since it may still contain quantified propositional and predicate variables; the remainder of the proof is devoted to showing how those may be eliminated.

First of all, we can replace any quantified propositional variable p in B by its instances $\boldsymbol{t}$ and $\boldsymbol{f}$, so we need only eliminate the predicate variables. Next, given two models $\mathcal{M}_1 = (U, \ldots)$ and $\mathcal{M}'_1 = (U', \ldots)$ of L_1, we write $\mathcal{M}_1 \leq \mathcal{M}'_1$ if $\mathcal{M}_1$ is a substructure of $\mathcal{M}'_1$ in the usual sense, but for which $U = U'$. The relation $(\mathcal{M}_1, Q|\mathcal{M}_1) \leq (\mathcal{M}'_1, Q|\mathcal{M}'_1)$ is defined in the same way, so that when this holds, $Q|\mathcal{M}_1$ is the restriction to the predicates of $\mathcal{M}_1$ of $Q|\mathcal{M}'_1$, in accordance with the Locality Principle. Suppose both structures are models of A($\mathbf{Q}$); then by assumption, $Q|\mathcal{M}_1 = Q_U$ on the predicates in $\mathcal{M}_1$ and $Q|\mathcal{M}'_1 = Q_U$ on the predicates in $\mathcal{M}'_1$. Moreover both are equivalent to B on the respective classes of predicates. Hence, given $p_1, \ldots, p_n$ predicates in $\mathcal{M}_1$, $B(p_1, \ldots, p_n)$ holds in $\mathcal{M}_1$ if and only if it holds in $\mathcal{M}'_1$. In other words, B is invariant under $\leq$ extensions in the sense of Feferman [3].[5] It follows from Theorem 4.2, p. 47 of Feferman [3] that we can choose B to have quantifiers only over individuals; in addition, since we have a constant $\mathbf{t}^{(k)}$ of each propositional and predicate sort, we can take B to have no free variables other than $p_1, \ldots, p_n$. In other words, B is a first-order formula in the usual sense, with all quantified variables being of the individual sort, which defines Q_U in each $\mathcal{M}_1$ when restricted to the predicates of $\mathcal{M}_1$. Lifting B to $B\!\uparrow$ and $\mathcal{M}_1$ to the corresponding $\mathcal{M}_2$ gives, finally, the desired result.

Discussion and Questions.

1. **Comparison with Zucker [18].** Zucker considers formal quantifiers Q at every finite type level, within which he deals with first order quantifiers (i.e. those at type level 2 whose arguments are predicates of type level 1) as a special case. (The case of higher types uses different arguments with both positive and negative results.) He denotes by S_c ('c' for 'classical') the set [of operations] $\{\wedge, \neg, t, \forall\}$. By way of comparison, it is worth quoting him at some length as to his aims (the italics in the following are Zucker's):

> We are looking for an argument of the following form: given a proposed new 'logical operation' (say a quantifier), show that it is explicitly definable in terms of S_c. ... Now what does it mean, to "propose a new quantifier Q for inclusion in the language?" Clearly, a symbol 'Q' by itself is useless: a *meaning* must be given along with it. ... In

[5]These are called outer extensions in Feferman [3], but in the case at hand they are just ordinary extensions with one sort fixed (or "stationary" in the language of that paper), namely the sort of individuals.

> fact a symbol 'Q' is never given alone: it is generally given together with a set of *axioms and/or inference rules*, proposed for incorporation in a logical calculus. Now we [make] the following *basic assumption*:
>
> > For Q to be considered as a *logical* constant, its 'meaning' must be *completely contained* in these axioms and inference rules.
>
> In other words, it is quite *inadequate* to propose a quantifier Q for incorporation in the calculus as a logical constant, by giving its meaning in set theory, say (e.g., "there exist uncountably many"), and also axioms which are merely *consistent* with this meaning. The meaning of Q must be *completely determined* by the axioms (and rules) for it: they must carry the *whole weight* of the meaning, so to speak; the meaning must not be *imposed from outside* (by, e.g., a set-theoretical definition), for then we merely have a 'mathematical' or 'set-theoretical' quantifier, not a *logical* one. ... Our basic assumption, then, gives a necessary condition for a proposed new constant to be considered as purely logical. We re-state it as a principle of *implicit definability*:
>
> > (ID) A logical constant must be defined implicitly by its axioms and inference rules.
>
> Hence in order to prove the adequacy of S_c, it will be sufficient to show that any constant which is *implicitly definable* (by its axioms and rules) is also *explicitly* definable from S_c." (Zucker [18], pp. 518–19)

There follow three notes (ibid.). The first is that (ID) is only proposed as a necessary (but not necessarily sufficient) condition for logicality. The second is that the inference rules for the new constant need not be of the natural deduction kind. Third, it is assumed that the status of the members of S_c as logical constants is not in doubt.

As noted in the introductory discussion above, one essential difference I have with Zucker is that I regard the meaning of a quantifier to be provided from the outside so to speak, i.e., to be given in model theoretic terms prior to the consideration of any rules of inference that may be in accord with it. For me, the significance of the condition ID is to specify completely its role as an inferential agent.

2. **What is a necessary and sufficient condition for logicality?** Taking for granted that the standard operations of FOL are logical, it is at first sight plausible that any quantifier defined in terms of them should also be considered logical. However, in a personal discussion following a presentation of this material,[6] Lauri Hella questioned this. He pointed out that many mathematical notions considered as Lindström quantifiers that would not ordinarily be considered logical are definable in FOL. For example, we can thus define what it is for (U, P) to be a group, where P is a ternary relation for the product relation of the group. Note that the relation of equality is used in this definition, and it is a matter of some contention whether equality is a logical notion (cf. Quine [15], pp. 61ff and Feferman [4], p. 44). We can side-step that issue by considering the definition in FOL without equality of all (U, P, E) where E is a congruence

[6] At the 2011 Workshop on Logical Constants in Ljubljana referred to in Footnote 7.

relation with respect to a product relation P under which the structure forms a group. The collection Q of all such (U, P, E) would still not ordinarily be considered to be a logical quantifier. In any case, that is the reason why the combined semantical and inferential criterion considered here is only proposed as a necessary condition. In order for such to be tightened to a necessary and sufficient condition, we would have to be explicit about what would constitute axioms and rules of inference for a quantifier Q that determine it uniquely. The work of Zucker and Tragesser [19], pp. 10–15 is a start on that for a formulation in natural deduction terms, but that needs to be generalized and, if possible, simplified.

3. **Extension to countable admissible languages.** It is shown in Feferman [2, 3] that the results from those articles needed for the proof here of the Main Theorem hold equally well for the sublanguages L_A of the language with countably long conjunctions and disjunctions and ordinary quantification, and for which A is an admissible set. Thus one should expect that the Main Theorem carries over directly to those languages. But now there is a new question that ought to be considered, namely whether all infinitary *propositional* operations that satisfy a necessary criterion for logicality similar to the one taken here, are definable in L_A.
4. **Are there analogous results for intuitionistic FOL?** There are several possible options to consider for the semantics of general quantifiers looked at constructively: the most familiar ones are the (so-called BHK) interpretation in terms of primitive notions of construction and constructive proof, realizability interpretations, inferential semantics, and Kripke models. It is an open question how Lindström quantifiers might be treated with respect to either of the first two of these. As to the third, one would take the work of Zucker and Tragesser [19] as a point of departure as suggested at the end of item 2 above; it is shown there (under certain natural hypotheses about the forms of inferences) that every formal quantifier given by introduction rules is equivalent to one definable in intuitionistic FOL. Finally, given any Lindström quantifier Q viewed classically, one can extend its semantics to arbitrary Kripke structures $(W, \preccurlyeq, \langle U_w : w \in W\rangle, \ldots)$ for which $w \preccurlyeq v$ implies $U_w \subseteq U_v$, by taking a formula $Q\underline{x}_1 \ldots \underline{x}_n(\psi_1(\underline{x}_1, \underline{y}), \ldots, \psi_n(\underline{x}_n, \underline{y}))$ to be satisfied by $\underline{b}$ in $(U_w, \ldots)$ just in case $(U_v, p_1, \ldots, p_n)$ is in Q for each $v \succcurlyeq w$, where P_i is the set of all k_i-tuples a_i in U_v such that $\psi_i(\underline{a}_i, \underline{b})$ is true at v. Then the definition of forcing works as usual. Since Kripke semantics reduces to classical semantics on worlds W having a single element, the Main Theorem can be applied to show that any Lindström quantifier on Kripke structures dealt with in this way and that satisfies the criterion considered here is definable in classical FOL. This leaves open whether some more intrinsic version of the Main Theorem holds for Kripke structures and intuitionistic FOL.[7]

---

[7]The main body of material for this article was first presented for a talk at the ESSLLI Workshop on Logical Constants, Ljubljana, Aug. 9, 2011. A second presentation was made on May 23, 2012 at a conference at CUNY in honor of Sergei Artemov, on the occasion of his 60th birthday.

Acknowledgements I would like to thank Jeremy Avigad, Denis Bonnay, Fredrik Engström, Lauri Hella, Michael Rathjen, Dag Westerståhl and Jeffery Zucker for their useful comments on a draft of this article.

References

1. Barwise, J., and S. Feferman. (eds.). 1985. *Model-theoretic logics*, Perspectives in mathematical logic series. New York: Springer.
2. Feferman, S. 1968. *Lectures on proof theory. Proceedings of the summer school in logic, Leeds*. Lecture notes in mathematics, vol. 70, 1–107.
3. Feferman, S. 1968. Persistent and invariant formulas for outer extensions. *Compositio Mathematica* 20: 29–52.
4. Feferman, S. 1999. Logic, logics and logicism. *Notre Dame Journal of Formal Logic* 40: 31–54.
5. Feferman, S. 2010. Set-theoretical invariance criteria for logicality. *Notre Dame Journal of Formal Logic* 51: 3–20.
6. Gentzen, G. 1935. Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift* 39: 176–210; 405–431. (English translation in Gentzen [7], 68–131).
7. Gentzen, G. 1969. *The collected papers of Gerhard Gentzen*, ed. M.E. Szabo. Amsterdam: North-Holland.
8. Gómez-Torrente, M. 2002. The problem of logical constants. *Bulletin of Symbolic Logic* 8: 1–37.
9. Hacking, I. 1997. What is logic? *Journal of Philosophy* 76: 285–319.
10. Keisler, J. 1970. Logic with the quantifier "there exist uncountably many". *Annals of Mathematical Logic* 1: 1–93.
11. Lindström, P. 1966. First order predicate logic with generalized quantifiers. *Theoria* 32: 186–195.
12. McGee, V. 1996. Logical operations. *Journal of Philosophical Logic* 25: 567–580.
13. Peters, S., and D. Westerståhl. 2006. *Quantifiers in language and logic*. Oxford: Clarendon Press.
14. Prawitz, D. 1965. *Natural deduction. A proof-theoretical study*; 2nd edn. (2006). Mineola: Dover Publications.
15. Quine, W.V.O. 1986. *The philosophy of logic*, 2nd ed. Cambridge: Harvard University Press.
16. Sher, G. 1991. *The bounds of logic*. Cambridge: MIT.
17. Tarski, A. 1986. What are logical notions? *History and Philosophy of Logic* 7: 143–154.
18. Zucker, J.I. 1978. The adequacy problem for classical logic. *Journal of Philosophical Logic* 7: 517–535.
19. Zucker, J.I., and R.S. Tragesser. 1978. The adequacy problem for inferential logic. *Journal of Philosophical Logic* 7: 501–516.

Chapter 3
Implicit Definitions, Second-Order Quantifiers, and the Robustness of the Logical Operators

Arnold Koslow

Abstract We use a modified version of E.Beth's concept of implicit definitions to show that all the usual logical operators as well as the first and second order quantifiers are implicitly defined—and for essentially the same reason that involves an account of the logical operators using a concept of filter conditions. An "inferential" proposal is then suggested for a Gentzen-like account as a necessary condition for the familiar logical operators. We then explore the question of whether our proposal can also be taken as a sufficient condition. To this end, we discuss whether other operators, like a truth operator, the counterfactual conditional, the identity, and the modal operators are also logical operators. The paper closes with a brief discussion of what is called the robustness of the logical operators: What happens to the logical operators when there is a shift from one logical structure to another which extends it, and what happens when there is a shift from one structure to one in which it is homomorphically embedded.

3.1 The Problem

We want to say something about the quantifiers of Second-order logic by placing them in a broader context. A great deal of interest has focused on what these quantifiers quantify over—sets, predicates, functions, or properties, for example. There have even been questions as to whether Second-order logic is set theory in disguise, and perhaps not even logic at all. I do not want to settle any of those issues here. What I do want to do is to make a case for the Second-order quantifiers as logical operators, in exactly the same way that the first-order quantifiers as well as negation, conditional, conjunction, and disjunction are logical operators. The cogency of this claim will depend of course on how we think of the logical operators. I also want to show that the grounds for their qualifying as logical operators can be settled affirmatively, without having to settle the ontological issues

A. Koslow (✉)
Philosophy Program, The Graduate Center, CUNY, 365 Fifth Ave., Rm. 7113,
New York, NY 10016, USA
e-mail: akoslow@mindspring.com

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_3

surrounding these quantifiers. The situation here is parallel to that of regarding the first-order quantifiers as logical operators, without having to settle what the first-order quantifiers quantify.

3.2 Implicit Definitions

There's a fairly long tradition in mathematics for thinking of mathematical theories, especially geometry, as characterizing their basic concepts implicitly.[1] Hilbert, in his Geometry[2] said in his famous exchange with Frege that his axioms defined the concepts of his theory. P. Bernays however in several explanations of the revolutionary significance of Hilbert's Geometry described the axioms as providing an implicit definition of the concepts. What he, or anyone around that time meant by "implicit definition" is not at all obvious, and to my knowledge, never explained.

It was Evert Beth, however, who offered two equivalent explanations of that term—one syntactic, the other model-theoretic, and proved that in a broad range of cases, if a theory provided an implicit definition of a term of a theory then that term also had an explicit definition provided by the other terms of that theory. According to Beth, if a theory implicitly defines a term—say a predicate, relation, function, or some constant, then that narrows down the possibilities to essentially just one. If a theory uses a term like "electron", why couldn't any other term be used instead, perhaps with another meaning, without changing the truth overall. Sometimes it might be claimed, you can't substitute any old term. It has to be one with the same meaning or same reference. The possibilities are limited. What Beth did was to provide a condition that limited the possibilities of substitution by appealing to the theory itself, and not to the meaning or reference of the terms in question.

Here is Beth's syntactic version of his notion of implicit definition (the original version)[3]:

Let $T(P)$ be a theory with a predicate P . Let $T(P^*)$ be the result of replacing all occurrences of "P" in $T(P)$, by the predicate P^* which does not occur in $T(P)$.

(EB) $T(P) \wedge T(P^*)$ implies $(\forall x)[P(x) \leftrightarrow P^*(x)]$.

Beth then proved a dramatic theorem: for a broad range of theories, if a term A is implicitly defined by some theory, then it is also explicitly defined by that theory. That is, the theory implies an if and only if statement (usually quantified) connecting A with a term B that is defined using the terms of the theory other than A. It is important to note that the biconditional of A with B is not to be thought of as a definition of A because it has some kind of necessity. It is not like the simple case of the definition of "brother" as "male sibling". The truth of the biconditional is

[1] Nagel [22].

[2] Hilbert [12].

[3] Beth [3, 4].

guaranteed by the truth of the two versions of the theory from which it logically follows. Consequently, whatever necessity the biconditional might have is derivative from whatever necessity the theory may have. It does not have it on its own. There have been several different proofs of Beth's definability theorem, for a broad range of theories. The theories under consideration could be physical, mathematical, or logical. We shall be concerned here with just the logical theories, both first and second-order and we will not be concerned with Beth's famous theorem that connects the concepts of implicit with explicit definition. Instead we shall focus on Beth's concept of implicit definition itself.

There are cases of course where a theory implicitly defines some of its terms, and Beth's remarkable theorem sheds light on those cases. However there are important cases where the theory implicitly defines (some of) its terms, but there are no explicit definitions to be had. For example, the classical sentential calculus formulated with just conjunction and negation provides, as we shall see, implicit definitions for those operators, but no explicit definitions of either operator in terms of the other. Of course if the classical sentential calculus is formulated with just one logical operator (either the Nicod or the Sheffer stroke, say), then even if one of them is implicitly defined, there is no other operator which can be used to give an explicit definition of it. The case of the intuitionistic sentential calculus is interesting. All the sentential logical operators (negation, conjunction, disjunction, and the conditional) are implicitly defined, as we shall see, but none of them is explicitly definable in terms of the others.

In the case of mathematical theories, there are important cases where certain important terms are not even Beth implicitly defined. One shouldn't assume that even when a mathematical theory is well known and important, that it provides implicit definitions of its basic terms. For example, despite what Hilbert and Bernays said, the key concepts of "point", "line", and "plane" of Hilbert's formulation of Euclidean plane geometry are not implicitly defined by the theory. In fact Hilbert, when pressed by Frege to define "point" presumably in terms of the remaining concepts of his formulation of the theory, replied that those concepts were defined (not implicitly defined) by the axioms. We know from what he said elsewhere about tables, chairs, and beer mugs, that the axioms did not define those concepts implicitly.

Here is an important example that has given so much trouble to defenders of a structuralist account of mathematical theories (cf. the insightful recent paper of Stewart Shapiro[4] on this example). Let *Comp*[*i*] be complex number theory with "*i*" the square root of -1. Suppose it true. Now everywhere in the theory *Comp*[*i*], replace all occurrences of "*i*" by "$-i$". The result is the theory *Comp*[$-i$]. It follows that "*i*" is not implicitly defined by complex number theory. If it were then by (EB), *Comp*[*i*] and *Comp*[$-i$] would imply that $i = -i$. But that's impossible.

In light of these examples, it becomes obvious that the notion of an implicit definition has an importance of its own, even when it does not figure as a successful

[4] Shapiro [25].

application of the Beth Definability Theorem. And it becomes important to avoid thinking that the mere occurrence of a term in a theory automatically insures that it is implicitly defined by that theory. There are cases where is it, and cases where it isn't.

If the theory has constants, predicates, and functions, then Beth's account covers three possibilities:

1. If the theory T contains an occurrence of the predicate letter "P", then replacing it everywhere by a predicate letter P^* (not occurring in T), yields the theory $T(P^*)$, and $T(P) \wedge T(P^*)$ implies $(x)[P(x) \leftrightarrow P^*(x)]$. Obvious adjustments if there is a relation R in the theory T. Then we say that T implicitly defines the predicate P or the relation R.
2. If the theory T has an occurrence of a constant letter "c", then replacing it everywhere by a constant letter c^* (not occurring in T) yields the theory $T(c^*)$, and $T(c) \wedge T(c^*)$ implies $c = c^*$.
3. If the theory T has an occurrence of a function letter "f", then replacing it everywhere by another function letter "f^*" (not occurring in T) yields the theory $T(f^*)$, and $T(f) \wedge T(f^*)$ implies $(x)[f(x) = f^*(x)]$. Obvious adjustments should be made for functions of several arguments. Then we say that the theory T implicitly defines the function f.

In the present paper, I want to use Beth's idea of implicit definition to discuss the implicit definability of the various logical operators. There is however, a technical problem. The usual discussions of Beth's notion are confined to terms that are either predicates, relations, functions, or constants of the appropriate language in which the theory is expressed. We would like to extend his idea to the logical connectives, such as conditionals, negations, conjunctions and disjunctions, as well as the universal and existential operators, whether they are first or second-order. It's clear that the usual sentential connectives, and the quantifiers are not included among the predicates, and functions for which Beth's definition was initially designed.

Here is an example of how we think the notion of implicit definition should be expressed for the case of connectives and quantifiers, while keeping to the heart of Beth's concept. Suppose then that $\mathcal{L}(\wedge)$ is some logical system (we shall sharpen what that means presently) with a connective for conjunction, and that $\mathcal{L}(\wedge^*)$ is the logical system that results from the uniform substitution of $\wedge^*$ for $\wedge$ in $\mathcal{L}(\wedge)$. Then, in the case for connectives we would say that $\wedge$ is implicitly defined by $\mathcal{L}(\wedge)$ if and only if

$$\mathcal{L}(\wedge)\&\mathcal{L}(\wedge^*) \text{ implies that } A \wedge B \Leftrightarrow A \wedge^* B \text{ for all instances of the schema.}$$

In requiring that each of he instances of the unstarred and starred operators mutually imply each other, we follow the way that Belnap handled what he called the "uniqueness" of the connective for conjunction.[5] This leaves us with a problem of how to extend the definition to include the universal and existential quantifiers,

[5]Belnap [2]. I understand Belnap as advocating the kind of schematic expression of the condition for implicit definitions that we made explicit. There may be reason to think otherwise. L. Humberstone [13, p. 267] reports that in correspondence Belnap suggested that the kind of

both first and second-order. They are not counted as connectives.[6] I shall for present purposes follow Mostowski's characterization of them, and assume that they are operators, or functions that act on predicates and relations, yielding a predicate (or relation) of lower arity. In order to treat the question of the implicit definability of these and other operators, I shall use the notion of an implication structure to simplify the following discussion.

By an implication structure $\mathcal{I} = \langle S, \Rightarrow \rangle$, we mean any non-empty set S, together with an implication relation, $\Rightarrow$, on it. And by an implication relation we shall mean any relation that satisfies essentially the Gentzen [8] structural conditions.[7] That is

1. **Refexivity**: $A \Rightarrow A$, for all A in S.
2. **Projection**: $A_1, A_2, \ldots, A_n \Rightarrow A_k$, for any $k = 1, \ldots, n$.
3. **Simplification** (sometimes called Contraction): If $A_1, A_1, A_2, \ldots, A_n \Rightarrow B$, then $A_1, A_2, \ldots, A_n \Rightarrow B$, for all A_i and B in S.
4. **Permutation**: If $A_1, A_2, \ldots, An \Rightarrow B$, then $A_{f(1)}, A_{f(2)}, \ldots, A_{f(n)} \Rightarrow B$, for any permutation f of $\{1, 2, \ldots, n\}$.
5. **Dilution**: If $A_1, A_2, \ldots, A_n \Rightarrow B$, then $A_1, A_2, \ldots, A_n, C \Rightarrow B$, for all A_i, B, and C in S.
6. **Cut**: If $A_1, A_2, \ldots, A_n \Rightarrow B$, and B, $B_1, B_2, \ldots, B_m \Rightarrow C$, then $A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m \Rightarrow C$.

We take these conditions as an axiomatic characterization of the notion of an implication relation. Deducibility relations, and semantic consequence relations are of course included. They are special important cases, but they are not the full story.[8] Noticeably absent in (1)–(6) is any reliance on logical operators. Negation, conjunction, disjunction, and the universal and existential quantification can now be defined for any implication structure, solely in terms of whatever implication relation is in place. The use of implication structures rather than some very specific system of logic will enable us to discuss the issue of the implicit definition of the logical operators in a very general setting, without the distraction of irrelevant special features of particular formulations of particular logical systems.

uniqueness he had in mind might also be thought of as implicit definition as long as some kind of second order (propositional quantification) is included. Here it has been Beth who was the inspiration for using his notion of implicit definition carried over to connectives and operators rather than just relations, functions, and constants. One shouldn't however overlook the important and unjustly neglected paper in J.Harris [10].

[6]Church [5, p. 39] for example, places them in the special category of operators.

[7]Gentzen [9], and Hertz [11].

[8]Cf. Koslow [16].

Key to the discussion then is the "definition" of all of the usual logical operators in terms of the implication relation that is specified in an implication structure. We turn then to a uniform method of defining all of the usual logical operators, and some other operators as well.[9]

3.3 Logical Operators on Implication Structures

Let $\mathcal{I} = \langle S, \Rightarrow \rangle$ be an implication structure. We will characterize the logical operators of negation, conjunction, disjunction, the conditional and first-order quantification in a way that reflects Gentzen's introduction and elimination conditions, but does so in a way that reveals an underlying uniformity which we shall make evident. To that group we shall then add second order quantification as a genuine cousin of the other logical operators.[10]

(i) **Conjunction**. The conjunction operator, $\wedge$, on an implication structure is a function of two arguments, such that for any A and B in S, $(A \wedge B)$ is a conjunction of them if and only if the following two conditions are satisfied:

1. $(A \wedge B) \Rightarrow A$ as well as B, and
2. $(A \wedge B)$ (if it exists), is the weakest member of the structure to satisfy this condition. That is, if T is any member of the structure, and if $T \Rightarrow A$, as well as B, then $T \Rightarrow A \wedge B$.

(ii) **Negation**. The negation operator, $\sim$, on a structure is a function of a single argument such that for any member A of the structure, its negation $\sim A$, (if it exists) is a member of the structure that satisfies two conditions:

1. A together with its negation imply all the members of the structure, and
2. For any T in the structure, if T together with A implies everything in the structure, then $T \Rightarrow \sim A$. The negation of A is the weakest member of the structure to satisfy the first condition.

(iii) **Disjunction.** The disjunction operator, $\vee$, on an implication structure is a function of two arguments, such that for any A and B in S, $(A \vee B)$ is a disjunction of them if and only if the following two conditions are satisfied:

1. For any T in the structure, if $A \Rightarrow T$ and $B \Rightarrow T$, then (if it exists) $(A \vee B) \Rightarrow T$, and

[9]The account is broadly "inferential", originating with P. Hertz, and G. Gentzen, and including N. Belnap [2]. D. Prawitz [23], M. Dummett [6], J.I. Zucker [27], R.S. Tragesser [27, 28] V. McGee [18, 19], and most recently S. Feferman [7]. I cannot specify the bibliographic detail as to dates and pagination, as the paper is included in the present. Profound apologies for the omission of the many significant semantic and proof-theoretical recent work that also deserve close study.

[10]For an extended study of the following paired conditions on these logical operators, cf. Koslow [16].

2. It is the weakest member of the structure to satisfy the first condition. That is, for any U in S, if [for all T in S, [if $A \Rightarrow T$ and $B \Rightarrow T$, then (if it exists) $(A \vee B) \Rightarrow T$], then $U \Rightarrow A \vee B$].

(iv) **Conditional**. The conditional operator, $\rightarrow$, on an implication structure is a function of two arguments such that for any A and B in S,

(a) $A, (A \rightarrow B) \Rightarrow B$, and
(b) It is the weakest member of the structure to satisfy the first condition. That is , for any T in S, if $A, T \Rightarrow B$, then $T \Rightarrow (A \rightarrow B)$.[11]

(v) **Universal First-order quantifiers**. A full account of the structuralist treatment of first order quantification would be beyond present space limitations It has however been treated at length elsewhere (Koslow [16]). The short story is that the preceding notion of an implication structure has to be extended to include a non-empty set of "objects" O. The notion of a predicate of the structure can then be defined (using a modification of the Tarski-Mostowski approach) as a mapping (with finite support) from infinite sequences of members of O to the members of the set S of the structure (Tarski),[12] and the notion of a universal (existential) quantifier as an operator that maps predicates to predicates and lowers the arity by one (Mostowski).[13] So, in the specification of this extended implication structure, we need to include the set Pr of the predicates of the structure. If $P(x)$ is any predicate (of arity 1) then the universal quantification of P is $(\forall x)Px$ (or $(\exists x)Px$, in the case of the existential quantification of P), then quantification of P would be a sentence of the structure. If the arity of P is greater than one, then its universal or existential quantification would be a predicate of the implication structure , but of lower arity. It is not difficult to prove that for any predicate P in the set of predicates Pr,

1. $(\forall x)Px \Rightarrow Pa$, for all a in O, and
2. If T is any sentence in the structure, and $T \Rightarrow Pa$ for all a in O, then $T \Rightarrow (\forall x)Px$.

[11] Although these conditions suffice for showing that the logical operators are implicitly defined by the paired conditions, more is needed to obtain the full story: All these conditions need to be stated in a more general way: let Γ stand for any finite (possibly empty) sequence of members of the implication structure. Then, for example, the two modified (parameterized) conditions for conjunction are: for any Γ, A, and B, (1) $\Gamma, A, (A \rightarrow B) \Rightarrow B$, and (2) It is the weakest member of the structure to satisfy the first condition. That is, for any T in S, if $\Gamma, A, T \Rightarrow B$, then $\Gamma, T \Rightarrow (A \rightarrow B)$. With this adjustment, if follows that for any A and B in a structure, if their conjunction exists in the structure, then $A, B \Rightarrow A \wedge B$. The fuller account is in Koslow [16], ch. 15.

[12] Tarski [26, p. 191].

[13] Mostowski, A. On a Generalization of Quantifiers. In Mostowski [21].

That is, (1) requires that the universal quantification of a predicate implies all the instances of that predicate, and (2) says that the universal quantification of P is the weakest sentence to imply all its instances.

We can now state two observations that are important for the remainder of this paper. The first is that there is an important connection between our first and second conditions for each of the logical operators, with the corresponding Gentzen *elimination* and *introduction* conditions that are paired with each of the logical connectives. The connection between Gentzen's version, and our reformulation is close, with the possible exception of the negation operator. For example, in the case for conjunction, our condition, (1) $(A \wedge B) \Rightarrow A$ as well as B, corresponds to Gentzen's elimination condition for conjunction, and our second condition, (2) if $T \Rightarrow A$, as well as B, then $T \Rightarrow A \wedge B$, corresponds to Gentzen's introduction condition for conjunction (Gentzen [9]). In general, for each of the logical operators familiarly counted as such, Gentzen's elimination condition corresponds to our first kind of condition, and his introduction conditions correspond to our second condition (the operator is the weakest satisfying the first condition).

The second observation redescribes the conditions we provided, in a slightly more formal way which will enable us to isolate a uniform requirement that covers all of the familiar logical operators in one fell swoop. Moreover, it is that redescription which will enable us to provide very simple proofs that all the familiar logical operators are implicitly defined by the preceding descriptions of them.

3.4 Logical Operators and Filter Conditions

Let $\mathcal{I} = \langle S, \Rightarrow \rangle$ be an implication structure. We shall say that $\mathcal{F}$ is a filter on the structure if and only if

1. If it is a non-empty subset of S, and
2. Any member of S that implies a member of $\mathcal{F}$ is also a member of $\mathcal{F}$.

$\mathcal{F}$ is a proper filter if and only if it is a proper subset of S. Filters are sets closed from above, and are the duals of sets that are closed under implication—i.e. Tarskian theories. If we look over the preceding characterizations of the logical operators of negation, conditionals, disjunctions, conjunctions and the first-order quantifiers, we see that in each case two conditions are provided. A filter condition is specified that is characteristic for the particular operator, and uses only the implication relation of the structure. The first condition says that the value of an operator—say $\phi(A)$, is in the filter. This corresponds to what Gentzen called the *elimination condition*. The second condition says that the value $\phi(A)$ is the weakest member of that filter. This condition corresponds to what Gentzen called the *introduction condition* of the operator.

Before we illustrate how all these logical operators are implicitly defined it is useful to have some examples. Here is how, for example, negation, conjunction, and universal quantification (First-order) look according to this filter account:

Negation: The filter condition is $\mathcal{F}^N_A(T) : [A, T \Rightarrow C]$, for all C, and

1. $\sim A$ satisfies the filter condition. i.e. $\mathcal{F}^N_A(T)(\sim A)$.
2. The weakest condition: For any sentence T in the structure, [If $A, T \Rightarrow C$, for all C, then $T \Rightarrow \sim A$].

Conjunction: The filter condition is $\mathcal{F}^C_{A,B}(T) : [(T \Rightarrow A$, and $T \Rightarrow B]$, and

1. $A \wedge B$ satisfies the filter condition. i.e. $\mathcal{F}^C_{A,B}(A \wedge B)$.
2. The weakest condition: For any sentence T in the structure, [If $T \Rightarrow A$, and $T \Rightarrow B$, then $T \Rightarrow (A \wedge B)$].

Universal Quantification (First-order): The filter condition $\mathcal{F}^U_P(T) : [T \Rightarrow Pa$, for all a in O and

1. $(\forall x)Px$ satisfies the filter condition, i.e. $\mathcal{F}^U_P((\forall x)Px)$
2. The weakest condition: If T is any sentence in the structure, and [$T \Rightarrow Pa$ for all a in O, then $T \Rightarrow (\forall x)Px$].

These paired conditions for each of the logical operators we shall call the "Filter theory" of the logical operators. Thus the universal quantifier (and the existential) and the other logical operators of negation, the conditional, disjunction and conjunction uniformly satisfy the filter theory. We leave that simple fact to the reader.

We could stop here with this characterization of the operators. There is however a feature which we believe the filters associated with the familiar logical operators have and which deserves emphasis. The feature can be easily explained by consideration of the negation operator. The associated filter is $[A, T \Rightarrow C]$, for all C in the implication structure $\mathcal{I} = \langle S, \Rightarrow \rangle$. This is schematic in A. If you replace it by some particular item in the structure, then the negation of that item is the weakest member of the filter. For some members of S, the filter might be all of S, and for other members it might be a proper subset of S. So the feature that we have in mind can be expressed this way (the proper filter property):

(PFP) For any (familiar) logical operator on a structure, some of its associated filters will be proper filters.[14]

To that extent the first-order quantifiers are logical operators in the same way, and if operators of negation, conjunction etc. are thought of as logical operators, then so too for the first-order quantifiers.

We shall see that the second-order quantifiers also satisfy filter theory, and therefore qualify as logical operators if their sentential cousins count as logical operators. Being second-order doesn't make them set-theoretical, or part of some specific mathematical theory, unless we are prepared to admit that there are parts of certain specific mathematical theories (e.g. set theory for example) that are logical

[14]We set to one side the relatively simple proofs that each of the familiar logical operators satisfies the proper filter property.

operators—*only in a different setting*. We treat them exactly in the way that we treat the other operators that are considered to be logical operators, regardless of what they quantify over. In this way they are logical operators in the same way that first-order quantifiers are logical operators, -independently of what they quantify over.

Universal Quantification (Second-Order): For a consideration of second-order quantification, we need to extend our notion of implication structures further to express quantification over predicates and functions—i.e. Full Second-order logic.

Usually accounts of Second-order quantification try to cleave as closely as possible to First-order accounts, and our version is no different in that matter. Ordinarily, different types of variables are used to indicate quantification over objects, predicates, and functions. Here, for present purposes, we shall use X, Y, Z, ... and not mark the difference between the various universal quantifiers by extra clauses with specially dedicated notation. We now extend the implication structures so that beyond S, the set of objects O and the set of predicates Pr, we add the set PPr of predicates of the predicates (E. Mendelson [20, pp. 376–89], V. McGee [18, pp. 54–78, esp. p. 63], [19]). The second-order quantification of any predicate ϕ in PPr is described this way.[15]

For any ϕ in PPr, and all predicates P in Pr, and any sentence T in S , the filter condition for Second-order universal quantifications $\mathcal{F}^{U2}\phi(T)$ is: [$T \Rightarrow \phi(P)$, for all predicates P in Pr], and

1. $(\forall X)\phi(X)$ satisfies the filter condition. That is $(\forall X)\phi(X) \Rightarrow \phi(P)$, for all predicates P in Pr,
2. The weakest condition: If for any sentence T, $T \Rightarrow \phi(P)$, for all predicates P in Pr, then $T \Rightarrow (\forall X)\phi(X)$. That is, $(\forall X)\phi(X)$ is the weakest member of the filter.

At this point, it is worthwhile noting that not every operator on an implication structure is a logical operator. The operator "Tonk" is a good example.

3.5 "Tonk" Is Not a Logical Operator

We have thus far characterized logical operators by a uniform type of schema that makes use of a characteristic filter condition associated with each operator. The famous "tonk" operator was put forth by Prior,[16] to essentially wreck the program

---

[15]We have generally used versions equivalent to the treatment of the second- order quantifiers to be found in Mendelson [20, pp. 376–89] and Vann McGee [18, pp. 54–78, esp. p. 63]. Not everything however is listed (no inclusion of the comprehension schema), but enough to be able to prove implicit definability.

[16]Prior [24].

of defining logical operators in terms of their role in implication. It is worth showing that although "tonk" can be a sentential operator on an implication structure—a particularly nasty one, it is not a logical operator.

Let $\mathfrak{I} = \langle S, \Rightarrow \rangle$ be an implication structure with an implication relation $\Rightarrow$ on it, and suppose too that there is a binary sentential operator τ (tonk) on the structure such that if any A and B are in the structure, so too is $A\tau B$, and the following two conditions are satisfied:

$\tau 1.$ $A \Rightarrow A\tau B$, and
$\tau 2.$ $A\tau B \Rightarrow B$.

It follows immediately that for any A and B, that $A \Rightarrow B$, and so any members of the structure are equivalent ($A \Leftrightarrow B$). It also follows that the tonk operator is implicitly defined by ($\tau 1$) and ($\tau 2$). That is, if we replace all occurrences of τ by τ^*, resulting in ($\tau^* 1$) and ($\tau^* 2$), then we have for any A and B, that it follows from ($\tau 1$), ($\tau 2$), ($\tau^* 1$) and ($\tau^* 2$), that $A\tau B \Leftrightarrow A\tau^* B$. Nevertheless, although tonk is implicitly defined by the implicational story about it, tonk is not a logical operator.

Here's the argument: Suppose that there is a non-empty proper filter for tonk, i.e. $\mathcal{F}\tau(T)$. Let us suppose that some A is in it, i.e. $\mathcal{F}\tau(A)$ since the filter is not empty. Let B be any element in the structure. We know that $B \Rightarrow A$. Therefore B is in the filter. So the filter contains every member of the structure, and therefore it is not a proper filter. So there is no filter condition for the tonk operator. Consequently, it is not a logical operator.

3.6 Filter Conditions and Implicit Definitions

We now can offer very simple proofs that each of these operators is implicitly defined by their associated filter-theories.

Negation: Suppose (1) $F^N_A(\sim A)$, and (2) If for any T, if $A, T \Rightarrow C$, for all C, then $T \Rightarrow \sim A$, and (3) $\mathcal{F}^N_A(T)(\sim * A)$, and (4) If for any T, if $A, T \Rightarrow C$, for all C, then $T \Rightarrow \sim * A$. From (1) and (4) we have $\sim A \Rightarrow \sim * A$. From (2) and (3) we have $\sim * A \Rightarrow \sim A$. Therefore, $\sim A \Leftrightarrow \sim * A$.

Conjunction: Suppose (1) $\mathcal{F}^C_{A,B}(A \wedge B)$ and (2) If for any T, if $T \Rightarrow A$ and $T \Rightarrow B$, then $T \Rightarrow A \wedge B$, and (3) $\mathcal{F}^C_{A,B}(A \wedge^* B)$ and (4) If for any T, if $T \Rightarrow A$ and $T \Rightarrow B$, then $T \Rightarrow A \wedge^* B$. From (1) and (4) we have $A \wedge B \Rightarrow A \wedge^* B$. From (2) and (3) we have $A \wedge^* B \Rightarrow A \wedge B$. Therefore $A \wedge B \Leftrightarrow A \wedge^* B$.

Universal Quantification (First-order): Suppose (1) $\mathcal{F}^U_P((\forall x)Px)$ and (2) if for any T, $T \Rightarrow P(a)$ for all a in O, then $T \Rightarrow (\forall x)Px$, and (3) $\mathcal{F}^U_P((\forall^* x)Px)$ and (4) If for any T, if $T \Rightarrow P(a)$ for all a in O, then $T \Rightarrow (\forall^* x)Px$. From (1) and (4) we have $(\forall x)Px \Rightarrow (\forall^* x)Px$. From (2) and (3) we have $(\forall * x)Px \Rightarrow (\forall x)Px$. Therefore $(\forall x)Px \Leftrightarrow (\forall^* x)Px$.

Universal quantification (Second-order): It should be clear then that $(\forall X)\phi(X)$ is implicitly defined by (1) and (2). For suppose that there is another quantifier,

$(\forall^* X)$ that satisfies (1) and (2). Since $(\forall X)\phi(X)$, and $(\forall^* X)\phi(X)$ are both the weakest member of the filter to imply $\forall\phi(P)$, for all predicates P in Pr, they are equivalent —i.e. $(\forall X)\phi(X) \Leftrightarrow (\forall X^*)\phi(X)$.

Thus far we have made a case for regarding the logical operators on a structure in terms of the implicationally defined filters on implication structures. Our use of filters is a gloss on Gentzen's idea that introduction and elimination conditions constitute a definition of each of the logical operators. As they stand, we agree with Gentzen when he said that in the ultimate analysis they provide definitions; only we think that they provide implicit definitions rather than explicit ones.

Up to now we have discussed the filter conditions as a necessary condition for logical operators. What about requiring that the filter conditions are sufficient conditions? That opens the possibility that the filter conditions may over generate. But it is worth while I think, for the present, to explore that possibility.

We shall briefly describe several interesting cases of operators on implication structures: (1) Truth (*Tr*), (2) the counterfactual conditional, $\Box\!\!\rightarrow$ of some system such as D. Lewis' system VC, and (3) some modal operators, and (4) the Identity operator. On our account of the logical operators, we think the truth operator "*Tr*" is a logical operator, the counterfactual conditional of VC and the identity operator are possible cases, and the modal operators generally are not.

3.7 The Truth Operator Is a Logical Operator

Recall that the truth operator "*Tr*" (as against the truth predicate) is a function mapping sentences to sentences of say an implication structure, such that the schemas $Tr(A) \Rightarrow A$, and $A \Rightarrow Tr(A)$, hold for all sentences A in the structure.

To show that, "*Tr*" is a logical operator we need first to specify the filter for it. We take that to be $F_A^W(X) : [X \Rightarrow A]$, for all sentences X in the structure. Since $Tr(A) \Rightarrow A$, we have that (1) $Tr(A)$ satisfies the filter condition—i.e. $F_A^W(Tr(A))$ holds. This is our version of the elimination condition for "*Tr*". Our version of the introduction condition for *Tr* is given by the weakest condition: (2) If X is any sentence of the structure, and $X \Rightarrow A$, then $X \Rightarrow Tr(A)$. In particular, take X to be A. Since $A \Rightarrow A$ It follows then that $A \Rightarrow Tr(A)$ (and in fact, this in turn implies (2)).

Clearly, "*Tr*" is an operator on the structure that satisfies the filter theory. Assuming that the filter conditions are sufficient for being a logical operator, it follows that the truth operator is a logical operator.[17]

Moreover, since it is a logical operator, it follows that it is implicitly defined by the filter theory for the truth operator. That result however may seem to be moot. Tarski remarked that if truth is implicitly defined by a theory Γ, then it ought to be categorical in the sense (now no longer in use) that $\Gamma(Tr)\&\Gamma(Tr^*)$ implies that

[17]We leave to one side the proofs that the truth operator, as well as the counterfactual conditional and the identity operators to be discussed, all satisfy the proper filter property (PRP).

$(\forall x)(Tr(x) \leftrightarrow Tr^*(x))$—that is, the extensions of the predicates "*Tr*" and "Tr^*" are the same.[18] Tarski regarded it as a defect of his theory that it did not yield that result. His definition of "categoricity for *Tr*" seems to be the one that Beth would later use for the implicit definition of any predicate of a theory. There is no conflict however, with our present claim. We used the truth operator, not the truth predicate, and as we noted above , our notion of implicit definition is meant to cover schemata involving operators and connectives, for which the notion of sameness of extension is inappropriate.

3.8 Is the Counterfactual Conditional a Logical Operator?

In particular one might wonder whether D. Lewis' counterfactual system VC implicitly defines "$\Box\!\!\rightarrow$".[19] Here's a possible affirmative answer. Drop from VC the axiom which guarantees centering: for any A and B, $A \wedge B \Rightarrow A \;\Box\!\!\rightarrow B$, and add the conjunction of

(VC1) $\sim(B \Rightarrow A \;\Box\!\!\rightarrow B)$,
(VC2) $\sim(\sim A \Rightarrow A \;\Box\!\!\rightarrow B)$, and
(VC3) $A, A \;\Box\!\!\rightarrow B \Rightarrow B$.

(VC1) and (VC2) insure that the counterfactual conditional is different from the material conditional, and (VC3) insures that the counterfactual conditional satisfies modus ponens (in fact this is already one of the axioms of VC). We add it here because we want to consider the conjunction of (VC1)–(VC3) with "T" replacing all the occurrences of $A \;\Box\!\!\rightarrow B$. Let $\mathcal{F}^{CF}_{A,B}(T)$ be the conjunction of

(i) $\sim(B \Rightarrow T)$,
(ii) $\sim(\sim A \Rightarrow T)$, and
(iii) $A, T \Rightarrow B$.

It is easy to see that $\mathcal{F}^{CF}_{A,B}(T)$ is a filter. In the usual way, we take the elimination condition for the counterfactual conditional $A \;\Box\!\!\rightarrow B$ to be (1) $\mathcal{F}^{CF}_{A,B}(A \;\Box\!\!\rightarrow B)$, and the introduction condition for the counterfactual conditional $A \;\Box\!\!\rightarrow B$ to be (2): If $\mathcal{F}^{CF}_{A,B}(T)$, then $T \Rightarrow A \;\Box\!\!\rightarrow B$. That is, the counterfactual is the weakest member of the filter.

It follows that if $\Box\!\!\rightarrow^*$ replaces every occurrence of $\Box\!\!\rightarrow$ in our slightly amended version of VC, that if we have both the filter theory for $(A \;\Box\!\!\rightarrow B)$ and the filter theory for $(A \;\Box\!\!\rightarrow^* B)$, then we also have that $(A \;\Box\!\!\rightarrow B) \Leftrightarrow (A \;\Box\!\!\rightarrow^* B)$. Thus the counterfactual conditional of our modified VC, is implicitly defined by the elimination and introduction conditions. It therefore qualifies as a logical operator.

---

[18]Cf. the exchange on this point in Ketland [14, pp. 69–94]; [15, pp. 1075–79], and Bays [1, 1061–73].

[19]Lewis [17, pp. 132–33].

There are two caveats for this result. Although we did drop the centering condition from Lewis' original version of VC, a close version of the centering axiom is restored by use of the elimination and introduction conditions that we provided. That is, by (2), if for any A and B such that B does not imply T, and the negation of A does not imply T, and T implies $A \rightarrow B$, then $T \Rightarrow A \;\Box\!\!\rightarrow B$. In particular, take T to be $A \wedge B$. Then $A \wedge B \Rightarrow A \;\Box\!\!\rightarrow B$. What results is a qualified form of Lewis' centering axiom. The conclusion, provided this modified version of Lewis' system VC is acceptable, shows that the counterfactual, on this account, is a logical operator. The second caveat is this: although our adjustment (the dropping of the centering axiom in favor of a weakened version and the addition of (VC1) and (VC2)) to Lewis' VC appears to be minimal, it is an open question whether this adjusted version is a viable theory of counterfactuals.

3.9 Modal Operators

The modal operators in all their glorious variety have always been an important kind of operator in logic, ever since Aristotle wrote of them. So the question is whether they are logical operators. The answer is, surprisingly, not all that clear. We believe that some are not, and that it is an open question whether any of them are.

Consider the modal systems T, K4, and S4, for example. They can be recast as implication structures, where the box operator maps the sentences of the structure to itself, and satisfies the conditions (1) $\Box A$ implies A, (2) $\Box A$ implies $\Box\Box A$, and (3), the conjunction of (1) and (2) respectively. Next, consider two modal operators $\Box$ and $\Box^*$ on a finite implication structure whose set S is $\{A, B, C, D\}$, and whose implication relation is given by the following conditions: A implies B, and implies C. B implies D, and C implies D cf. Fig. 3.1.

These implications are only one way. Let $\Box$ be the operator which maps A to A, C to A, B to B, and D to D. On the same structure, let $\Box^*$ be the mapping of A to A, B to A, C to C, and D to D. It is easy to check that negation is classical on this structure, and for any thesis of this structure (the bottom member D which is implied by every member of the structure), the two operators acting on the thesis D, box(D), and box*(D), are again theses. Both box, and box* are T-modals). However neither of these modal operators implies the other—box of B doesn't imply box* of B, and box* of C doesn't imply box of C. Therefore the T-modal condition (1) does not implicitly define that modal. Moreover, the same operators also satisfy the

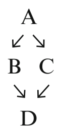


Fig. 3.1 A four-element implication structure

conditions for being a K4 modal and an S4 modal, but the conditions (2), and (the conjunction of (1), and (2) do not implicitly define K4 and S4 modals. Consequently, we know that some modals systems do not implicitly define their respective modal operators.

We have argued that every logical operator on a structure is implicitly defined by that structure, so that these modal operators are not logical operators. Consequently, not every modal of every modal system is implicitly defined by that system.

However, *some* modal systems do implicitly define their associated modal operators. L. Humberstone [13, pp. 603–05], has a very nice way of extending some modal systems so as to obtain a modal system which does implicitly define the system's modal operator. I have, in the following, slightly modified his proof: Take any modal system in which the modal is normal and the box of any thesis is a thesis. Add on these two further conditions:

3. $\Box(A \to A)$ is a thesis, and
4. $\Box(A \leftrightarrow B), C(A) \Rightarrow C(B)$, where $C(B)$ differs from $C(A)$ in having one or more occurrences of A replaced by B.

To show that this modal is implicitly defined, we assume that there is another modal $\Box^*$ which satisfies the same conditions, and then show that $\Box(A) \Leftrightarrow \Box^*(A)$ follows from the starred and unstarred versions of (3) and (4). So, assume also that

3*. $\Box^*(A \to A)$ is a thesis, and
4*. $\Box^*(A \leftrightarrow B), C(A) \Rightarrow C(B)$, where $C(B)$ differs from $C(A)$ in having one or more occurrences of A replaced by B.

Given these four assumptions and that all these modals are normal, it follows that $\Box(A) \Leftrightarrow \Box^*(A)$. Modifying Humberstone's proof slightly, note that $\Box(A)$ is equivalent to $\Box(A \leftrightarrow [R \to R])$, that $\Box^*(A \leftrightarrow [R \to R])$ is equivalent to $\Box^*A$, that $\Box^*(A \to A)$ is a thesis, and take it to be $C(A)$, then by (3) we have $\Box(A \leftrightarrow [R \to R]), \Box^*(A \to A) \Rightarrow \Box^*(A \leftrightarrow [R \to R])$, so that $\Box(A) \Rightarrow \Box^*(A)$. The converse is proved similarly.[20] Of course the Humberstone observation doesn't show that all modals are implicitly defined, but some are when special features are added to them.

So the boxes of some modal systems are implicitly defined by those systems, and some are not. The question at hand however is not whether modal operators are implicitly defined by their associated systems; it is the question of whether they are logical operators. The fact that some modals are not implicitly defined, carries with it the conclusion that some modal operators are not logical operators. There is still the open question of whether there are some modal operators that are logical operators.

[20]It is clear that the heavy lifting in this argument is supplied by (4) (and its mate (4*)). These conditions, as Humberstone notes are the modal analogs of the substitutivity condition on identity, (5) $t = u, C(t) \Rightarrow C(u)$, for which is well known to imply that any two relations satisfying it mutually imply each other.

3.10 The Identity Operator

The logical operators considered thus far were operators on implication structures simple, or extended, and they vary in type. The simple ones were given by a non-empty sets together with an implication relation on it. We considered "sentential" operators, mappings of S to S, or the Cartesian product of $S \times S$ to S. Next there were the extended structures, which consisted of a non-empty set S, a set of objects, and a set of predicates Pr. We discussed the logical operators such as the universal (and existential) quantifiers on the extended structures, first and second-order. These were operators that mapped predicates to predicates, and included the special case when quantifiers acted on predicates and had the sentences of S as values. We now turn to a logical operator, the identity operator that is a mapping from pairs of objects in O to the members of S.

Let e and e' be any members of O, and P be any predicate of Pr. Let $I(e, e')$ be the member of S, to be read as "e is identical with e′". We think of the identity operator on I, as that which assigns to the pair (e, e') of objects in O, the sentence $I(e, e')$. Let the filter associated with the identity operator I, be $\mathcal{F}^I_{e,e'}(T)$: $[T \Rightarrow (\forall P)(P(e)$ if and only if $P(e')]$. This is a filter. Then we have the two filter conditions for identity: (1) $\mathcal{F}^I_{e,e'}(I(e, e'))$. That is, the value of the identity operator for the pair (e, e') satisfies the filter condition. Therefore, if $I(e, e')$ then $P(e)$ if and only if $P(e')$ for all predicates P of the structure. This is just the principle of the indiscernibility of identicals. The second condition (2) (the weakest) requires that $I(e, e')$ is the weakest element to satisfy the filter condition, i.e. if $\mathcal{F}^I_{e,e'}(T)$, then $T \Rightarrow I(e, e')$. However $(\forall P)[P(e')$ if and only if $P(e)]$ satisfies the filter condition, so if $(\forall P)[P(e')$ if and only if $P(e)]$, then $I(e, e')$. This is of course just the identity of indiscernibles. In our version we have the nice conclusion that the elimination condition for the identity operator is the indiscernibility of identicals, and the introduction condition for the identity operator is just the identity of indiscernibles. That is, the identity operator is a logical operator if and only if indiscernible A and B are identical and identical A and B are indiscernible. The two principles are a package deal and belong together.

3.11 Concluding Remarks: The Robustness of the Logical Operators

There is a feature of the logical operators, not much noticed, that can be explained easily by using implication structures. Consider two structures $\mathcal{I} = \langle S, \Rightarrow \rangle$, and $\mathcal{I}^* = \langle S^*, \Rightarrow^* \rangle$, where $\mathcal{I}^*$ is a conservative extension of $\mathcal{I}$. It is an extension, so that $S \subseteq S^*$, and $\Rightarrow \subseteq \Rightarrow^*$, and conservative, so that for all $A_1, \ldots, A_n$, and B in S, we have $A_1, \ldots, A_n \Rightarrow B$, if and only if $A_1, \ldots, A_n \Rightarrow^* B$. This is the case of one structure being a substructure of another. And in this case it is easy to show that if for example "$A \wedge B$" is the conjunction in $\mathcal{I}$ of A and B, and "$A \wedge^* B$" is the conjunction of A and B in the structure $\mathcal{I}^*$, then in $\mathcal{I}^*$,

$$(A \wedge B) \Leftrightarrow^* (A \wedge^* B).$$

Thus, even though the implication relation in each of the two structures may be different, and the conjunction of members of $\mathcal{I}$ may not be the same as the conjunction of those members in $\mathcal{I}^*$, nevertheless they will be equivalent in the extending structure. Thus, what is a conjunction in the first structure will continue to be a conjunction in the extending structure, and the two conjunctions will be equivalent in the extending structure. This is what I want to call the robustness of the conjunction operator, in the shift from one implication structure to another.

Similar remarks hold for the other logical operators, only in their case an additional assumption is needed in order to obtain the same result. Here's the proof for the negation operator: Assume that there are two implication structures, $\mathcal{I}$ and $\mathcal{I}^*$, that S is a subset of S^*, that the implication relation $\Rightarrow^*$ is a conservative extension of $\Rightarrow$, and (in addition) the set U is closed under the negation N^* operator of $\mathcal{I}^*$. (In general we shall say that a set U is closed under an operator ϕ if and only if for every A in U, $\phi(A)$ is equivalent to some C in U.) Then we have the result that for every A in S,

$$N(A) \Leftrightarrow^* N^*(A).$$

That is, in passing from the first implication structure to the second, the negation of A in the first structure and the negation of A in the second structure are equivalent according to the implication relation of the second structure.

Thus far, we know that robustness holds when passing from one implication structure to another that is a conservative extension of it. What is even more impressive however, is the situation in which one structure is not a substructure of another , but instead, is embeddable in the second by some homomorphism ϕ.

Consider the case then when the members of $\mathcal{I} = \langle S, \Rightarrow \rangle$ needn't overlap with the members of $\mathcal{I}^* = \langle S^*, \Rightarrow^* \rangle$ and the implication relation $\Rightarrow$ needn't be a subrelation of $\Rightarrow^*$, but there is some function ϕ mapping S to S^* such that

1. For all $A_1, \ldots, A_n$ and B in S, $A_1, \ldots, A_n \Rightarrow B$ if and only if $\phi(A_1), \ldots, \phi(A_n) \Rightarrow^* \phi(B)$.

In the case of conjunction for example, it is easily seen that we have a homomorphism theorem (noted by Warren Goldfarb, in private correspondence)[21]:

$$\phi(A \wedge B) \Leftrightarrow^* \phi(A) \wedge^* \phi(B),$$

where "$\wedge$" is the conjunction operator on $\mathcal{I}$, and "$\wedge^*$" is the conjunction operator on $\mathcal{I}^*$. Thus the image (under ϕ) of the conjunction in $\mathcal{I}$ of A with B, is equivalent in $\mathcal{I}^*$ to the conjunction in $\mathcal{I}^*$ of their images.

[21] Cf. Koslow [16, p. 389, fn. 8].

This indicates a persistence of the conjunction operator over shifts from one implication structure to a second, where the first structure is embeddable in the second, and they can differ substantially in their respective sets and implication relations. If you start with a conjunction in an embeddable structure, you will not lose the conjunctive character. In the structure that embeds the first one, you will end up with a conjunction—not necessarily the same one that you had, but another, which in the second structure is equivalent to what you once had.

This result for all the logical operators seems to be to be a remarkable feature of them that indicates their special importance. Its proof for the conjunction operator is remarkably easy. Here's a proof:

$A \wedge B \Rightarrow A$, and $A \wedge B \Rightarrow B$, so that $\phi(A \wedge B) \Rightarrow^* \phi(A)$ and $\phi(A \wedge B) \Rightarrow^* \phi(B)$. Therefore $\phi(A \wedge B) \Rightarrow^* \phi(A) \wedge^* \phi(B)$. Conversely, $A, B \Rightarrow A \wedge B$ so that $\phi(A), \phi(B) \Rightarrow^* \phi(A \wedge B)$. Consequently, $\phi(A \wedge B) \Leftrightarrow^* \phi(A) \wedge^* \phi(B)$.

The proof for the other operators is slightly more complicated. One needs not only a conservative extension, but the requirement that the embedded structure is closed under the negation operator of the embedding structure. Thus if the structure $\mathcal{I} = \langle S, \Rightarrow \rangle$ is embeddable in the structure $\mathcal{I}^* = \langle S^*, \Rightarrow^* \rangle$ by the homomorphism ϕ, and the set S is closed under the negation operator of the structure $\mathcal{I}^*$, it follows that for every A in S,

$$\phi(\sim A) \Leftrightarrow^* \sim^*(\phi(A)).$$

That is, for every A, the image of its negation in $\mathcal{I}$ is equivalent in $\mathcal{I}^*$ to the negation of its image. Thus even when one structure is not a substructure of another, but is embeddable in it, the image of the negation of A is equivalent in the embedding structure to the negation of the image of A—or as we described this carry-over on negations from one structure to another, the negation operator is robust. So too are the other operators we have described as logical.

References

1. Bays, T. 2009. Beth's theorem and deflationism. *Mind* 118: 1061–1073.
2. Belnap, N.D. 1962. Tonk, Plonk and Plink. *Analysis* 22(6): 130–134.
3. Beth, E.W. 1953. On Padoa's method in the theory of definition. *Indagationes Mathematicae* 15: 330–339.
4. Beth, E.W. 1964. *The foundations of mathematics*. Amsterdam: North-Holland. 1959, Harper Torch books: 290–293.
5. Church, A. 1956. *Introduction to mathematical logic I*. Princeton: Princeton University Press.
6. Dummett, M. 1973. The philosophical basis of intuitionistic logic. Logic colloquium, Bristol, ed. H.E. Rose and J.C. Shepherdson, 5–40; reprinted 1978. Truth and other enigmas. London: Duckworth, 215–247; reprinted 1983. Philosophy of mathematics, 2nd ed., ed. P. Benacerraf and H. Putnam, 97–129. Cambridge University Press.

7. Feferman, S. Forthcoming. Which quantifiers are logical? A combined semantical and inferential criterion. In *Quantifiers, quantifiers, and quantifiers: Themes in logic, metaphysics, and language*, ed. A. Torza.
8. Gentzen, G. 1932. über die Existenz unabhängiger Aimsysteme zu unendlicher Satzsysteme. *Mathematische Annalen* 107(3): 329–50.
9. Gentzen, G. 1934–1935. Untersuchungen über das logische schliessen. *Mathematische Zeitschrift* 39: 176–210, 405–43; reprinted 1969. Investigations into logical deduction. The collected papers of Gerhard Gentzen, ed. M.E. Szabo, 68–131. Amsterdam: North-Holland.
10. Harris, J. 1981. What's so logical about the logical axioms. *Studia Logica* 41: 159–171.
11. Hertz, P. 1929. Über Axiomensystem für beliebige Satzsysteme. *Mathematische Annalen* 101: 457–514.
12. Hilbert, D. 1899. Grundlagen der Geometrie; 1902. The foundations of geometry. (English Trans. E.J. Townsend Open Court.; 1999. 14th German edition B.G.Teubner Stuttgart, Leipzig).
13. Humberstone, L. 2011. *The connectives*. Cambridge: MIT.
14. Ketland, J. 1999. Deflationism and Tarski's paradise. *Mind* 108: 69–94.
15. Ketland, J. 2009. Beth's theorem and deflationism—reply to Bays. *Mind* 118(472): 1075–1079.
16. Koslow, A. 1992. *A structuralist theory of logic*. Cambridge/New York: Cambridge University Press.
17. Lewis, D. 1973. *Counterfactuals*. Cambridge: Harvard University Press.
18. McGee, Vann. 2000. Everything. In *Between logic and intuition essays in honor of Charles Parsons*, ed. G. Sher and R. Tieszen, 54–78. Cambridge: Cambridge University Press.
19. McGee, Vann. Forthcoming. 2015. The categoricity of logic. In *Foundations of logical consequence*, ed. C. Caret and O. Hjortland Thomassen. Oxford University Press.
20. Mendelson, E. 2010. *Introduction to mathematical logic*, 5th ed. Boca Raton: Chapman & Hal/CRC.
21. Mostowski, A. 1979. *Foundational studies, selected works*, vol. II. Amsterdam: North Holland.
22. Nagel, E. 1939. The formation of modern conceptions of formal logic in the development of geometry. *Osiris* 7: 142–224.
23. Prawitz, D. 1965. *Natural deduction. A proof-theoretical study*. Stockholm: Almqvist & Wiksell; 2006, 2nd ed. Mineola: Dover Publications.
24. Prior, A. 1961. The Runabout inference-ticket. *Analysis* 21(6): 124–128.
25. Shapiro, S. 2012. An "I" for an i: Singular terms, uniqueness, and reference. *The Review of Symbolic Logic* 5(3): 380–415.
26. Tarski, A. 1956. The concept of truth in formalized languages. In *Logic, semantics, metamathematics*, ed. Alfred Tarski. Oxford: Clarendon Press. (Trans. by J.H. Woodger; 1983. 2nd ed., edited, Hackett Publishing company, Indianapolis and introduced by John Corcoran).
27. Zucker, J.I. 1978. The adequacy problem for classical logic. *Journal of Philosophical Logic* 7: 517–535.
28. Zucker, J.I. and R.S. Tragesser. 1978. The adequacy problem for inferential logic. *Journal of Philosophical Logic* 7: 1501–1516.

Chapter 4
Quantifiers Are Logical Constants, but Only Ambiguously

Sun-Joo Shin

Abstract Why is it crucial to categorize quantifiers as logical constants? After clarifying the importance of the issue in a larger context of logical theories, the paper investigates the following question: Why do we encounter a more contentious debate on the logical constancy of quantifiers than in the case of sentential connectives? Starting from the intuitive and naive rationale and moving to more complicated arguments for the well-accepted view that quantifiers are logical constants, I identify two tiers of the meanings assigned to quantifiers: At the first-level the interpretation is changing, and at the meta-level a constant meaning is assigned. I claim the tension arises from this ambiguous nature of the quantifier-semantics and illustrate the effects of the tension both in relevant literature and in a non-classical logic where the interpretation of the universal quantifier includes the empty domain. The double features of the universal quantifier – varying at one level and constant at another level – could raise skepticism toward the debate on logical constants itself, and in turn, toward the Tarskian analysis of logical consequence whose success heavily relies on the clear-cut status of logical constants.

Are quantifiers logical constants? Many have said "yes,"[1] but for different reasons, and a few have said "no" or have been skeptical of a definite answer. The question involves two debated topics in the philosophy of logic – quantifiers and logical constants. One might easily say "It all depends on what we mean by 'logical constants'." This is not incorrect. The response acknowledges that the concept of

[1] "[I]t is generally agreed that signs for negation, conjunction, disjunction, conditionality, and the first-order quantifiers should count as logical constants, . . . " (MacFarlane [5]).

S.-J. Shin (✉)
Department of Philosophy, Yale University, 108 Connecticut Hall, Old Campus, 344 College Street, New Haven, CT 06511, USA
e-mail: sun-joo.shin@yale.edu

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_4

logical constants is quite tricky and sometimes controversial, while many might think we know what quantifiers are. Yes, we can identify a given symbol as a quantifier, since it is a syntactic entity after all. However, as we will see below, how this syntactic piece is assigned its semantics doubly confuses our question “Are quantifiers logical constants?”

The paper is not about logical constants in general. First of all, I do not propose any necessary and sufficient features of logical constants. Moreover, I do not intend to get into an even bigger picture and make a claim for a position on the spectrum running from the essential nature of logical constants to the skepticism about the logical constant debate itself. Instead of making any direct claim about logical constants, the paper will focus on the way quantifiers have been addressed in the discussions about logical constants. More specifically, by exploring the way quantifiers are handled in formal semantics, I identify interesting and important features of quantifiers that the existing literature has overlooked and uncover our ambiguous stance regarding their status as logical constants. Our discussions, I strongly suspect, will only confirm how murky our concept of logical constants is, and at the same time will provide us with a new angle on existing literature on the topic.

The first section presents the background story for logical constants. Not being a survey of different views on the issue, it is the setting which helps us to understand the origin of the topic and, hence, to realize what issues are at stake in the logical constant debate. I will reveal and highlight certain assumptions one needs to accept when he/she gets into the discourse of logical constants. In the second section, starting with the most uncontroversial view that sentential connectives are logical constants, we will examine their properties as logical constants. These properties are cited and applied in the third section when we take on quantifiers. Are we fixing the meaning of quantifiers as we are for connectives? That story is not that simple. When the story is unfolded, we will see the complexity and the myth involved in our logical constants talk, quantifier talk, and beyond. By pointing out the difficulty in categorizing quantifiers either as logical constants or as non-logical constants, I would like to show how fragile and unclear our grasp of logical constants is. Nonetheless as seen in our preliminary section, when one subscribes to the most prominent logical theory, that is, the Tarskian view, he/she is not in a position to abandon the logical constant project. Our quantifier talk might, I suspect and hope, provide an occasion to re-think the Tarskian project.

4.1 Preliminaries

Logic is the study of valid reasoning, and a logical theory, exploring theoretical issues surrounding validity, presents its own coherent picture for important questions like the following: What makes one argument valid and another not? Does it have something to do with the meanings of the words in an argument, but nothing else? Do we need to conquer the meaning of meanings in order to conquer validity? (If so, would it be a hopeless project from the beginning?) What is the relation among logical, necessary, and *a priori* truth? No relation, whatsoever? Any one of these questions is controversial enough to consume philosophers' lives generation after generation. For better or for worse, we agree that logical constants have been quite often at the center of these hotly debated topics.

Up-front I would like to say that controversies surrounding logical constants, some believe, are a symptom of a wrong direction taken by the entire logical validity/logical truth project.[2] Then, the next step is to ask the following question:

(Q) Can we talk about validity (and logical truth) without talking about logical constants?

Answers differ. Let me start our preliminary discussions with question (Q) in order to show how logical constants have been a major topic in some logical theories while logical constants do not get in the picture at all for some other theories. A contrast in ways to explain the validity/non-validity of arguments (1) and (2) nicely illustrates why some say "no" and some say "yes" to question (Q):

(1) Every man is mortal.
Socrates is a man.
———————————
Socrates is mortal.

(2) Every man is mortal.
Socrates is mortal.
———————————
Socrates is a man.

A desideratum: We need to convince the reader that in the case of (1), it is impossible that the premises are true while the conclusion is false, and in the case of (2), it is possible to have true premises and a false conclusion. In spite of the common goal, how to think about impossibility/possibility cannot be more different in answer (A) versus answer (B):

Answer (A):
Imagine a world where every man is mortal and Socrates is a man. Then, in that world, it must be the case that Socrates is mortal. Hence, (1) is valid. On the other hand, I can imagine a world where every man is mortal and Socrates is mortal, but Socrates is a dog. Hence, (2) is not valid.

Answer (B):
Validity mainly has something to do with meanings, but not with how the world looks. We do not have to know the meaning of every word in a given argument, but only certain kinds of words. In the case of (1), 'man' could mean dog or cat, 'Socrates' could refer to Plato, etc. As long as 'every' has the meaning as we know, the truth of the premises guarantees the truth of the conclusion. On the other hand, in argument (2), when we interpret 'man' as cat, our conclusion is false while the premises are true. Hence, (2) is not valid.

There are many ways to point out differences between (A) and (B), and let's focus on the nature of a counterexample each side provides to assure that (2) is not valid. Approach (A) presents a possible *world* as a counter example, say w, such that the premises are true and the conclusion is false in w. On the other hand, (B) introduces a possible *interpretation* as a counterexample, say I, according to which the conclusion turns out to be false and the premises are true in the actual world.

[2] Etchemendy [2, Ch. 9].

That is, (A) fixes the interpretations/meanings of words and varies over worlds and (B) does the opposite. Etchemendy identifies the crucial differences between these two approaches and names (A) *representational* and (B) *interpretational*.[3] It is not my intention to evaluate one over another, but to draw out the essence of each position.

The representational approach is quite intuitive since the modal nature of the impossibility/possibility (of the true premises and a false conclusion) for arguments (1) and (2) is directly transferred to an impossible/possible world scenario. We can easily see this method ends up equating logical truth with necessary truth. Some might welcome this result and some might not. Another big obstacle for the representational approach is how to systematize it. We have modal intuitions, and for simple arguments like (1) and (2) the intuition does the job. However, if logical validity and logical truth are built on our modal intuition, we would feel somewhat backwards or question-begging. Whether a sentence is logically true seems to be independent of our world-views. If a (formal) theory is given to us, we might be able to test out whether the theory fits (some of) our modal intuitions, but as time-honored related philosophical controversies show us, we cannot come up with a formal system which utilizes modal intuition only.[4] First of all, intuition differs from one person to another. Second, searching for all possible worlds is next to impossible! Third, the judgment of possible/impossible worlds assumes our understanding of the meanings of a sentence or words. Here we run into another cart-before-horse scene: A logical theory which incorporates semantics as an important component is supposed to explain the mechanism of the birth of the meaning of a complex unit out of the meanings of simpler units, and a representational theory assumes this miraculous process happens (which is true) and tells us whether there is a possible world where these sentences are true. The more we think about it the more likely approach (A) belongs to metaphysics rather than to logic. Recalling that our project is to explain logical validity/non-validity, we start seeing the shortcomings of approach (A).

Sure enough, in answer (B) we do not find (heavy) modal terminology. The stake seems to be meaning or interpretation. While answer (A) demands us to imagine all possible worlds, answer (B) asks us to consider all possible interpretations. Everything seems to be taken care of in terms of interpretations/meanings. Hence, the success of this method would mean a successful reduction of modality to non-modality. Any misgiving we raised above against (A) would disappear here. That is exactly what Tarski's celebrated analysis of logical consequence aimed for. Many have believed that Tarski's work fits the bill and that model theory, which grew out of Tarski's analysis, is the end of the story about validity and logical truth.

There has been an exception to this consensus, though. At the end of the twentieth century, Etchemendy presented a book-length argument to show that Tarski's project fails in modal-reduction (conceptually) and could be saved extensionally only

[3] Etchemendy [2, Chs. 2 & 4].

[4] This is one of the reasons why we have various modal systems.

thanks to other extra assumptions and the weakness of a first-order language. Not surprisingly, there has been a strong skepticism toward Etchemendy's criticism against Tarski's project. Again, I am not entertaining this on-going debate in this paper, but would like to get to the heart of Tarski's project since that will lead us to the main topic of the section – the birth of logical constants.

How does Tarski's idea propose to analyze away modality? He does not provoke any possible world talk here, but only different assignments for certain terms in given sentences:

> [W]e have the concept of the *satisfaction of a sentential function* by single objects or by a sequence of objects. ... The intuitive meaning of such phrases as: *John and Peter satisfy the condition 'X and Y are brothers', or the triple numbers, 2, 3, and 5 satisfies the equation x+y=z",* can give rise to no doubts. ... One of the concepts which can be defined in terms of the concept of satisfaction is the concept of *model*. ... Let L be any class of sentences. We replace all extra-logical constants which occur in the sentences belonging to L by corresponding variables, ... [W]e obtain class L' of sentential functions. An arbitrary sequence of objects which satisfies every sentential function of the class L' will be called a *model* or *realization of the class K of sentences.* ... In terms of these concepts we can define the concept of logical consequence as follows:
>
> *The sentence X follows* logically *from the sentences of the class K if and only if every model of the class K is also a model of the sentence X.* [5]

For a given sentence, say α, we replace "extra-logical constants" with variables and make its sentential function α'. We call a sequence of objects which satisfies α' a model of α. Instead of 'every possible world' in the representational approach, Tarski presents 'every model', which is every possible interpretation. To put it simply, Tarski's logical consequence test is performed by allowing various (re)interpretations of words in an argument.

However, clearly, not all of the words should get re-interpreted to get the right result. If so, no argument would be valid. In the case of argument (1) above, if 'every' is re-interpreted as 'some' and 'mortal' as 'rich,' we find a counter example with interpretations which make the two premises true but the conclusion false. We want to say that this is not a genuine counterexample, by blocking the re-interpretation of the word 'every.' On the other hand, if we take every word to be a logical constant, argument (2) would be valid, which is not correct. Hence, approach (B) could be successful *only if* we would realize that there are certain (not all) words that we are not allowed to reinterpret. These are logical constants. The existence of logical constants is a necessary component of Tarski's analysis of logical consequence.

It is important to note that logical constants are an issue only in the context of an interpretational, not of a representational, theory. In the case of the representational account the meanings of words are fixed, and we change worlds. Every word should have its constant meaning. In the case of the interpretational account, we need to rule out two extreme scenarios: Either all of the word meanings are fixed or none of them is fixed. If we allow none of the words in an argument to change its meaning,

[5] Tarski [10, pp. 416–17].

then many non-valid arguments would be judged to be valid (like argument (2) above). If we allow all of them to get re-interpreted, then no argument would be valid (except when a conclusion is the same as a premise). Here is a rationale for a selection of logical constants.

What is so special about 'every' as opposed to 'man,' 'mortal,' and 'Socrates' in the above arguments? Could we come up with necessary and sufficient features of logical constants? Or, is it a relative or a pragmatic decision to decide which vocabulary items are logical constants?[6] Without attempting to tackle any of these questions, I would like to start with the most simple and uncontroversial collection of logical constants, that is, sentential connectives.

For the rest of the paper, I assume the following: (i) Approach (A) and approach (B) are fundamentally different, and so we need to choose one or the other (possibly neither), but not both. (ii) Only in the case of the interpretational approach, logical constants have become an important issue. Hence, our discussions will focus on Tarskian-style interpretational analysis. It is time to examine closely how meaning-(non)assignment takes place in interpretational semantics.

4.2 Semantics of Connectives

Let's take up the simplest logic – sentential logic – to see how sentential connectives are categorized as logical constants. There are three kinds of basic syntactic entities in sentential languages – sentential symbols, sentential connectives, and parentheses. After formal syntax tells us which strings are grammatically acceptable, formal semantics assigns meanings to those grammatical strings, i.e. sentences. Examining how meanings are assigned to various kinds of vocabulary might shed light on why we think connectives are logical constants.

The goal of formal semantics is to define the logical consequence relation in a rigorous way, so we need to come up with a systematic algorithm to assign T or F to any given sentence. To rule out meaning-ambiguity, a mathematical tool – the function – is adopted so that one and only one meaning is attached to a given sentence. It is important to note that there are two kinds of semantic functions at work in the case of sentential logic. One is a semantic function, say v, which assigns meaning, that is, T or F, to each sentence symbol, and the other is a semantic function, say C, for each sentential connective which defines the meaning of a connective. First of all, the two functions, v and C, are different in their domains;

[6]Tarski himself was quite agnostic about the issue in his paper: "Perhaps it will be possible to find important objective arguments which will enable us to justify the traditional boundary between logical and extra-logical expressions. But I also consider it to be quite possible that investigations will bring no positive results in this direction, so that we shall be compelled to regard such concepts as 'logical consequence', 'analytical statement', and 'tautology' as relative concepts which must, on each occasion, be related to a definite, although in greater or less degree arbitrary, division of terms into logical and extra-logical." (Tarski [10, p. 420])

one is a set of sentential symbols and the other a set of n-tuples of truth values. More importantly, I would like to draw attention to a hierarchical difference between v and C: (i) Function v assigns T or F to each sentence symbol. (ii) Given function v, function C is defined in terms of semantic values assigned by the semantic function v. Hence, let me call semantic function v a first-level function and semantic function C a meta-function over function v.

Another major difference between v and C is illustrated by the following simple example: Do we know the meaning of a sentence symbol, say A_1? No, until semantic function v is given to us. The meaning could be either T or F. How about sentence $(A_1 \,\&\, A_1)$? No again, until semantic function v is given to us. An interesting point is that all we need to know for the meanings of both A_1 and $(A_1 \,\&\, A_1)$ is function v, which I call first-level, precisely because we know the semantic function for '&', say $C_{\&}$. That is, the meanings of sentential connectives are assumed to be constant throughout a sentential logical system. This is the essence of our practice where we take connectives to be logical constants. As we will see below, the meaning-constancy of connectives is implemented both in the semantic function for sentences and other logical concepts, i.e. logical consequence, logical truth, and logical equivalence.

With varying semantic functions v and fixed semantic functions C,[7] a truth-table provides a decisive algorithm for the meaning of any arbitrary sentence, say α. Suppose n sentence-type symbols occur in α. Then, 2^n many possible v-functions are enumerated, and for each row the computation of the meaning of α is mechanically performed without any additional piece of information when (or since) semantic functions for connectives in α are assumed. A truth-table visually exhibits two important aspects which distinguish two kinds of semantic functions. (i) Functions v and C are of different levels: Given a row of truth-values, that is, a v-function, C-functions are applied to the values of the given v-function. (ii) While v-functions are different from row to row, C-functions are constant throughout the entire truth-table. The semantic function for sentences being recursively defined as an extension of function v, we may compute, given v, the meaning of any arbitrary sentence. That is, having v-functions as basic functions for the recursion, an extension of v-functions, i.e. $\bar{v}$, is defined in terms of C-functions.

Let's see how v-functions and C-functions feature in the following important logical concepts:

Sentence α is a *logical consequence* of a set of sentences Γ if and only if
every v-function that makes every member of Γ true also makes α true.
α is a *logical truth* if and only if every v-function makes α true.
α is *logically equivalent* to β if and only if
every v-function that makes α true also makes β true, and vice versa.

Throughout the three definitions, we may make the following quite intriguing observations about the two kinds of semantic functions we have been talking about:

[7] I do not want to say 'constant semantic functions' to avoid a confusion with constant functions in general use, e.g. $g(\mathbf{x}) = n$.

(i) The three logical concepts are defined in terms of certain conditions holding for every v-function.
(ii) C-functions are not mentioned in any definition.

What ultimately matters for these important logical concepts? What decides the extension of these concepts? Item (i) tells us no specific v-function matters at all. Item (ii) is even more puzzling: Does it tell us C-functions do not matter?

On the contrary, C-functions are assumed in every definition since the truth of a sentence is defined recursively in terms of C-functions as we said before. It would be awkward to say 'some C-functions' or 'every C-function.' There is one and only one semantic function for each connective while there are uncountably many v-functions. The examination of the above definitions tells us that for these logical concepts what matters is C-functions, not v-functions. In this sense, sentential logic is a logic of connectives.

So far we have explored the features of sentential connectives to account for the rationale for calling them logical constants. Let us step back from our common practice and raise the following questions: How constant are they? What justifies their constant meanings? Is it possible to change C-functions? Yes. Then we are doing a non-classical logic. There are many kinds of non-classical logic, and changing C-functions forms the basis of one of them.[8]

The meaning of the conditional symbol has been a source of dissatisfaction with the way classical logic handles sentential connectives. It is the beginning of C. I. Lewis' strict implication proposal as opposed to material implication, which opened the door for modal concepts into a logical system. Realizing Lewis' strict implication cannot avoid a similar paradox with conditionals, some logicians decided to implement relevance between propositions P and Q when we talk about the implication from P to Q. Hence, the meaning of the conditional symbol is modified: Relevance logic has its own C-functions (which are different from classical C-functions) and they remain constant throughout their system. Note that one and the same sentence[9] would get different truth values between classical and relevance logic. That is, the meanings of logical constants are not constant from logic to logic. Therefore, what we mean by logical 'constants' is logic-relative: Given a logic, C-functions remain the same.

On the other hand, even though C-functions are logic-relative, we do not re-write the definitions of logical concepts from logic to logic. Note that the extensions of these concepts change, though. That is, in one logic, sentence α is a logical consequence of Γ while not in another logic. That does not mean we are changing the concept of logical consequence or logical truth *per se*. The reason why these concepts remain the same while their extensions change (thanks to changes in C-functions) is that the role of logical constants in the logical concepts has not changed from logic to logic. Many of us have believed that there has been a strong consensus

[8] In the case of many-valued logic, changing the range of v-functions takes place.

[9] I am talking about a sentence as a syntactic object, not a proposition.

about the robust role of logical constants in doing logic. This is one of the reasons why the topic of logical constants has enjoyed its special and unique status in the philosophy of logic. When we move to quantified logic, however, things become tricky and unstable. Is it because of quantifiers? Or, is it because we do not have as clear a concept of logical constants as we believe we have? We will revisit these questions at the end.

4.3 Semantics of Quantifiers

4.3.1 Domain of Discourse

Quantifiers have been classified as logical constants, with some exceptions.[10] In the previous section, we located a distinction between logical constants and other syntactic units in how meanings are assigned to them. So, let's examine how semantic entities are attached to first-order logic vocabulary and check whether the main features we identified for connectives as logical constants also hold for quantifiers.

In addition to parentheses and sentential connectives, first-order languages have the following list of vocabulary: quantifier(s), predicates, names (optional), variables, and function symbols (optional). Unlike sentential languages, these syntactic units require heterogeneous kinds of semantic entities. An object is assigned for a name, a set of n-tuples of objects for an n-place predicate, and an n-ary operation over an n-tuple of objects for an n-ary function symbol. Informally, a sentence Pc_1c_2 is true if and only if the ordered pair of an object denoted by c_1 and an object denoted by c_2 is a member of the set of ordered pairs denoted by predicate P. Let me call these heterogeneous semantic assignments "p-functions."[11]

How about quantifier $\forall$? A non-empty set of objects is assigned to it. Let me call it a "Q-function." One way to address the main question of the paper (i.e. whether quantifiers are logical constants) is to raise the following question: Are there fundamental differences between p-functions and a Q-function? In the previous section, we identified the following four semantic factors which distinguish the logical constants (i.e. sentential connectives) and the other units (i.e. sentential symbols). First, the semantic functions for logical constants (C-functions in the above) are meta-operations over the basic semantic functions (v-functions) which assign meanings for basic vocabulary (sentence symbols in the case of sentential logic). The second aspect, which has been more the focus in the literature, is that C-

[10] Enderton [1, pp. 69–70] puts quantifier symbol $\forall$ into the parameter group, along with predicate symbols, constant symbols, and function symbols. Etchemendy, who is against the logical constants debate itself, experiments both with fixed and various meanings for $\forall$. (Etchemendy [2, Chs. 5 & 8].)

[11] I call them "p-functions," meaning semantic functions for parameters, as opposed to for logical constants.

functions are fixed while v-functions are not. Third, the meaning-constancy of C-functions plays a crucial role for logical concepts (like logical consequence, logical truth, and logical equivalence): the extension of these concepts is determined by the C-functions. Fourth, when we change C-functions, we move from one logic to another.

Let's compare p-functions with a Q-function in terms of these four aspects. First, is there a hierarchical difference? Is a Q-function a meta-operation over p-functions? Obviously not. Nonetheless, there seems to be a level-difference between a Q-function versus p-functions: A Q-function delineates the boundary of the other kinds of semantic entities. That is, the ranges of p-functions are determined by the Q-function: Names are interpreted as objects within the set which is the value of the Q-function. Similar things are going on with the interpretations of predicates and function symbols. Let's recall that the interpretation of $\forall$ is the domain of discourse. Even though there is no meta-structure between the Q-function and p-functions (as we found between C-functions and v-functions in the case of sentential languages), we can see what Q-function assigns serves as the background frame within which p-functions are situated. One may say that the Q-function is more fundamental than the rest of the semantic functions.

How about the second aspect of logical constants, that is, their meanings being fixed? Is the Q-function fixed as C-functions are for connectives? The truth of a sentence is determined by the interpretations of its components, and semantic functions are doing the job. In order to know the truth-value of sentence $\forall xRxx$, we need to know the interpretations of both the quantifier $\forall$ and the predicate R. Is the interpretation of quantifier $\forall$ fixed throughout a given logic? The Q-function assigns a set of objects, say set A, to $\forall$, and p-function for R assigns a set of paired orders of objects; hence the interpretation of R is a subset of the set $A \times A$. The only constraint for classical logic is that set A should be non-empty, but by no means we are talking about a fixed set; hence, there is no fixed meaning. Recall that in the case of sentential logic as long as we get the truth-values of α and β, we may compute the truth-value of sentence $(\alpha \ \& \ \beta)$ since the C-function for the connective & is fixed throughout a given logical system. That is not the case with the sentence $\forall xRxx$. As we noted in the previous section, the fixed meanings of connectives seem to support our common practice and belief that they are logical constants. On the other hand, the interpretation of $\forall$ is changing from one set to another (as long as it is not the empty set) since the domain of discourse is changing from discourse to discourse. I suspect this is the main reason why Enderton classifies $\forall$ as a parameter as opposed to a logical symbol.[12] Then, how can we justify the claim that quantifiers are logical constants? Why do many of us consider quantifiers to be logical constants? Keeping this question in mind, let's move to the third feature of logical constants on our list.

[12] Our logical constants are subsets of Enderton's logical symbols. Logical symbols are either those whose interpretations are fixed (i.e. connectives and the identity symbol) or those which are not interpreted (i.e. parentheses, commas, and variables); Enderton [1, pp. 69–70].

The next item on our checklist is the role logical constants play in logical consequence, logical truth, and logical equivalence. In the case of sentential logic, no semantic function other than C-functions matters in deciding on the extensions of these logical concepts. Given C-functions, we may decide whether α is logical consequence of Γ, whether α is logically true, or α and β are logically equivalent. No specific v-function is needed, since those logical concepts demand certain conditions hold for 'every' v-function. Does the Q-function in first-order logic play a similar role as C-functions do in the case of sentential logic? If so, we could define the logical concepts we examined before in the following way:

Sentence α is a *logical consequence* of a set of sentences Γ if and only if every p-function that makes every member of Γ true also makes α true.
α is a *logical truth* if and only if every p-function makes α true.
α is *logically equivalent* to β if and only if every p-function that makes α true also makes β true, and vice versa.

An easy counter example: Suppose a Q-function interprets $\forall$ as set {Tom, Mary}. Then, for a unary predicate R, there are four possible p-functions. They are:

$p_1(R) = \emptyset$.
$p_2(R)$ = { Tom }.
$p_3(R)$ = { Mary }.
$p_4(R)$ = { Tom, Mary }.

Then, the following statements would be true according to the above (pseudo) definitions:

- $\forall xRx$ is a logical consequence of $\sim \forall \text{x}\ \forall y((Rx\&Ry) \rightarrow x = y)$.
- $\forall x\forall y((Rx\&Ry) \rightarrow x = y)\ \vee\ \forall xRx$ is logically true.
- $\forall xRx$ is logically equivalent to $\sim\forall x\forall y((Rx\&Ry) \rightarrow x = y)$.

The main reason why we are getting wrong results is that throughout the definitions we fixed the interpretation of a quantifier as the domain which has only two members. Obviously, we need to revise the definitions so that certain conditions hold for every Q-function as well as every p-function.[13]

Sentence α is a *logical consequence* of a set of sentences Γ if and only if every Q-function and every p-function that makes every member of Γ true also makes α true.
α is a *logical truth* if and only if every Q-function and every p-function make α true.
α is *logically equivalent* to β if and only if every Q-function and every p-function that makes α true also makes β true, and vice versa.

In these definitions, there would be no reason to make a distinction between Q-functions and p-functions. According to classical model theory, a model is defined as a set of both Q-functions and p-functions. Hence,

[13]Please note that in the revised definitions, $\forall xRx$ is not a logical consequence of $\sim \forall \text{x}\ \forall y((Rx\&Ry) \rightarrow x = y))$, $\forall x\forall y((Rx\&Ry) \rightarrow x = y) \vee \forall xRx$ is not logically true, and $\forall xRx$ is not logically equivalent to $\sim\forall x\forall y((Rx\&Ry) \rightarrow x = y)$.

Sentence α is a *logical consequence* of a set of sentences Γ if and only if every model that makes every member of Γ true also makes α true.
α is a *logical truth* if and only if every model makes α true.
α is *logically equivalent* to β if and only if every model that makes α true also makes β true, and vice versa.

We need a pause here: According to the above definitions, there seems to be no distinction between the interpretation of quantifiers and the interpretations of other non-logical symbols and neither of them should be fixed. None of them plays the same role as C-functions do in the case of sentential logic. Our question revisited: Why have we considered quantifiers to be logical constants?

Our last item is to check whether a change of Q-functions correlates with a change of logic. We have seen that relevance logic departs from classical logic by changing some C-functions.[14] As we have seen, Q-functions are not as fixed as C-functions are. On the other hand, there is one aspect of Q-functions related to this issue: classical logic gives a constraint for Q-functions, that is, the interpretation of $\forall$ cannot be the empty set. When we allow the empty set as an interpretation of $\forall$, we are engaging in a non-classical logic called free logic. Here there seems to be a similarity between C-functions and Q-functions. We will take up this aspect later in the paper.

The interpretations of connectives mark a clear difference from the interpretations of other symbols both in their level and constancy. Connectives are interpreted over the semantic values of components of a sentence. The interpretations of connectives are so constant that a change in their meanings amounts to a change of logic. Things are not as clear-cut in the case of quantifiers as they are for connectives. The interpretation of quantifiers sets up a boundary for the interpretations of other syntactic parameters. In that sense, Q-functions and p-functions seem to have a slight difference in their semantic categories, but not a first-level versus meta-level hierarchical difference. Moreover, when it comes down to the meaning-constancy issue, Q-functions do not behave at all as C-functions do. A domain (which is the interpretation of $\forall$) is not fixed, and hence we cannot tell whether a random sentence is true even when other symbols are interpreted. Here is a contrast between $(A_1\&A_2)$ versus $\forall xRxx$.[15] Logical consequence, logical truth, and logical equivalence – all of them demand the domain not be fixed! There is one bottom line, though: The domain cannot be empty in classical logic. Hence, here is a tiny hint of constancy. When we include the empty domain, we are shifting from one logic to another.

[14]It might be slightly wrong to say that changing the meaning of the conditional symbol is the only motivation for relevance logic, but this is related to different views of logical consequence.

[15]If we know the truth values of A_1 and A_2, we may compute the truth value of $(A_1\&A_2)$. On the other hand, even if we know the interpretation of the predicate R, we cannot tell whether the sentence $\forall xRxx$ is true unless we know the domain, i.e. the interpretation of the universal quantifier.

4.3.2 Looking for Constancy

So far our discussions suggest quantifiers do not behave like connectives in terms of semantic assignments. If the universal quantifier denotes different sets from one model to another (which is the case), we cannot say its meaning is fixed. Then, why has the universal quantifier been considered as a logical constant? *In what sense* do many of us believe that the meaning of the universal quantifier is fixed? I would like to go through three different levels of responses, two in this subsection and the third one in the next subsection.

Let me start with a very simple and naive intuition behind our common belief that quantifiers are logical constants and call it the 'natural language argument.'

Every man is mortal.
Socrates is a man.
———————————
Socrates is mortal.

To check the validity of the argument, we could change the meanings or references of 'mortal,' 'man,' or 'Socrates,' but not 'every.' If we changed the meaning of 'every' to the meaning of 'some,' this argument would not be valid. The sentence "Grass is green or grass is not green" is logically true as long as we fix the meanings of 'or' and 'not.' At this intuitive level, we can easily see why connectives and quantifiers have been grouped together as logical constants.

Let's examine how the intuition behind the natural language argument is transferred into a first-order language so that $\forall$ may be justified as a logical constant. However, our examinations of the formal semantics of connectives and quantifiers concluded that there is a big difference between these two kinds of vocabulary in how constant their assigned meanings are. For example, we do not know the truth of the sentence "Everybody is a sophomore" until we know the domain of the discourse, even when we know who are sophomores. On the other hand, we know the truth of the sentence "Grass is green and snow is white" as long as we know the truth-values of "grass is green" and "snow is white." Using the terminology of (classical) first-order semantics, the interpretation of the universal quantifier, which is a non-empty set, is not constant. So, when the domain has only single object, the sentence $\forall x \forall y\ x = y$ is true, and otherwise it is false. Then, in the above simple argument in what sense do we assume the meaning of 'every' is fixed, and hence, a logical constant? Obviously, if we change 'every' to 'some' or 'most,' the argument is not valid anymore. That does not mean that the extension of 'every' is fixed, but *something about* 'every' is fixed.

Where does that constancy appear in first-order semantics? Let's carefully compare the semantic clauses of '&' and '$\forall$':

$(\alpha\ \&\ \beta)$ is satisfied by truth function v iff
  α is satisfied by v <u>and</u> β is satisfied by v.
$\forall x \phi(x)$ is satisfied by a model iff
  for <u>every</u> object in the domain, say a, $\phi(x/\text{a})$ is satisfied by the model.

As the natural language argument says, if we replace 'every' with 'some' in the above semantic clause, we do not get the right result. Let me call it the 'semantic clause argument.' The semantic clause for ∀ also seems to be very similar to the semantic clause for &. First of all, those clauses are not changing from one model to another; that is, things are fixed. The way ∀ is interpreted as 'every' is analogous to the way the semantic clauses are set up for connectives, e.g. ∼, &, ∨, etc. Like connectives, the universal quantifier is interpreted in a meta-language, and the meanings of their interpretations, say "it is not the case," "and," "or," and "every," are understood at a meta-level as if they are fixed.[16]

However, the similarity is so striking and obvious that we have overlooked the following interesting difference between ∀ and &, I claim. Let me make a contrast between a connective and a quantifier by rewriting the above semantic clauses in the following way[17]:

(α & β) is satisfied by a model, say $\mathfrak{I}$, if and only if
  α is satisfied by $\mathfrak{I}$ <u>and</u> β is satisfied by $\mathfrak{I}$.
$\forall x\phi(x)$ is satisfied by a model, say $\mathfrak{I}$, if and only if
  for <u>every</u> object in $\underline{\forall^{\mathfrak{I}}}$, say a, $\phi(x/\mathrm{a})$ is satisfied by $\mathfrak{I}$.

In the case of '&,' its interpretation/meaning "and" is given in a meta-language <u>once</u>. On the other hand, in the case of '∀,' interpretations take place <u>twice</u>: Once with 'every' (like 'and' for &) and also with $\forall^{\mathfrak{I}}$ as the domain of discourse. While the former meaning does not change from model to model, just like 'and', the latter interpretation is changing from model to model. Notation $\forall^{\mathfrak{I}}$ tells us its variance. Hence, the former aspect accounts for the view that ∀ is a logical constant, while the latter explicitly shows us that ∀ is non-logical vocabulary, i.e. a parameter, as Enderton classifies it.

Before delving into the matter further, I would like to address one immediate concern. Is it legitimate to bring in two interpretations for one and the same piece of vocabulary? Wouldn't it have caused an ambiguity? Interestingly, it does not cause an ambiguity, which turns out to be a mixed-blessing for our understanding of quantifiers. The two interpretations of ∀ come at two different levels – one as a set (varying from model to model at the first-level) and the other as 'every' (fixed at a meta-level). Hence, an ambiguity (as we usually worry about) does not take place. This might be one of the main reasons why no attention has been paid to the fact that the universal quantifier is interpreted *in two different ways* in its semantic clause.

[16]The flip side of the same story can be told in terms of inference rules. For each piece of vocabulary, we have inference rules and some think that the inference rules endow these logical vocabulary with its own meaning. If we take this proof-theoretic view, the point I am making here could be expressed in the following way: Both ∀ and other connectives are given their meanings when their own inference rules are introduced. (The relation between meanings and inference rules has been debated for decades – Do inference rules endow logical constants with meanings? Or, are inference rules justified in terms of the meanings which are given by semantic clauses of the logical constants? I have no intention to say anything about this directly, but let me emphasize that the point I am making here is consistent with both views.)

[17]I replace truth-function with model here.

Even though the two-tiered interpretation of ∀ has not caused a plain ambiguity, I suspect that the two levels of the ∀-semantics are a main source of the confusion and controversies we encounter in the literature on quantifiers, especially in the context of logical constants.

4.3.3 Dancing on Two Levels

Believing that quantifiers are logical constants as connectives are, many have attempted to find the common features between these two categories of vocabulary.[18] At the same time, realizing that quantifiers' constancy is different from connectives' constancy and requires more involved explanations, much literature has focused on quantifiers in the discussions of logical constants. Gila Sher's work on quantifiers is one of the most recent efforts in that direction. Sher, along other philosophers,[19] has defended Tarski's interpretational analysis of logical consequence. As we concluded in the first section, Tarski's logical consequence would not even get the right extension unless we get the right collection of logical constants. Unlike with connectives, quantifiers' interpretations do not easily convince us that they are constant. Being aware of the importance of and the difficulties involved with quantifiers as logical constants, Sher brings in more complicated and more sophisticated machinery than just appealing to 'every' in the semantic clause for a universal quantified formula cited above. Let me call this machinery the 'domain-size argument.'

Sharing the intuition that logical constants should be specific content-free, Sher wants to show they are *formal*. She defines the formality of logical constants in the following sprit of Mostowski: "A logical quantifier does not allow us to distinguish between different elements of [the universe]."[20] The next step is to extend this spirit to different domains of discourse but isomorphic to one another, and "being formal is being invariant under isomorphic structures."[21] Then, Sher draws our attention to the Homomorphism Theorem: If two models are isomorphic to each other, no first-order sentence can differentiate them. From that result, Sher concludes that ∀ is a logical constant. Since universal-quantified sentences turn out to be of the same truth-value in isomorphic models, the universal quantifier is not tied up with the content of a domain, but only with a formal character, mainly size, she argues.[22] Is this criterion convincing? As I explain below, Sher's contribution to the topic could

[18]Peacocke's [7] well-cited criterion is one prime example for this direction of effort.

[19]Refer to Gomez-Torrente [3] and Shapiro [8].

[20]Sher [9, p. 14], and refer to Mostowski [6] for more technical details.

[21]Sher [9, p. 53].

[22]MacFarlane [5] pushes the result further and makes things clearer: "Indeed, because cardinality is permutation-invariant, cardinality, every cardinality quantifier is included, . . . " (Section 5).

be better understood in the context of the two-tiered interpretations of ∀ I have identified and argued for.

There is a gap between the description of how quantified sentences are behaving among isomorphic models and the claim that quantifiers are logical constants. Leaving the logical constant agenda aside, the most immediate and innocent question we could raise about the Homomorphism Theorem is: Why do quantified sentences get the same semantic values in isomorphic models, even though their domains are different? It is not because quantifiers denote the same thing (which is not the case) or are constant (which is, at best, question-begging), but because our first-order languages are not expressive enough to reflect different contents of the domains. If so, the true import of the Homomorphism Theorem is the weakness of expressiveness of a first-order language. That is, if two models are isomorphic, there is no first-order sentence which distinguishes one from the other.

Things can get even shakier in light of the compactness theorem of first-order logic. Think about a non-standard model of arithmetic, say $\mathcal{M}'$ (where there is an object which is greater than any natural number), which is not isomorphic to a standard model $\mathcal{M}$. The Compactness Theorem takes us to the following proposition: Even though $\mathcal{M}$ and $\mathcal{M}'$ are not isomorphic to each other, for every first-order sentence, say α, α is true in $\mathcal{M}$ if and only if it is true in $\mathcal{M}'$. This result does not set us off to explore what that mysterious object is, but to convince us that our first-order arithmetic language is not expressive enough to pin down the world of numbers you and I normally conceive. So, being isomorphic (hence, each being of the same size in its domain) is a sufficient condition for the same semantic value of every quantified sentence, but not a necessary one. Then, we seem to be losing ground in tying up the constancy of a quantifier and the size of a domain. After all, it might be a test of the expressiveness of our language.

On the other hand, Sher's project could become more meaningful when the following issue is highlighted: Even when the domains of two models consist of different elements – hence, ∀ denotes different sets – universal quantified sentences turn out to be of the same semantic value as long as the models are isomorphic to each other. There must be *something* constant in the way ∀ plays a role in the semantics of quantified sentences. First of all, Sher realized that we need more or different explanations for ∀ as a logical constant than in the case of connectives, and mainly focused on quantifiers.[23] Let us recall that one level of interpretation of ∀ (say, the first-level interpretation) is a non-empty set. Given two different models,

[23]We should, therefore, not be surprised to run into the following comments on Sher's proposal: "Sher's proposal is somewhat awkward when understood as a general criterion for logical constancy since it has no straightforward application to sentential connectives." (Warmbrōd [11, p. 507]) Note her own definition of Tarskian logical constants: "*C* is a (Tarskian) *logical term* iff *C* is a truth-functional connective <u>or</u> *C* satisfies conditions of (A) to (E) above on logical constants [permutation-invariant especially among isomorphic structures]." (Sher [9, p. 56], and my underline.)

say $\mathfrak{I}$ and $\mathfrak{J}$, the interpretation of ∀ in $\mathfrak{I}$, say set I, and the interpretation of ∀ in $\mathfrak{J}$, say set J, are different from each other. Sher admits this varying aspect of the universal quantifier:

> Take the universal quantifier. In every model for a first-order logic the universal quantifier is interpreted as a singleton set (i.e., the set of the universe). But in a model with 10 elements it is a set of a set with 10 elements, whereas in a model with 9 elements it is a set of a set with 9 elements. Are these interpretations the same?
>
> I think that what distinguishes logical constants in Tarski's semantics is not the fact that their interpretation does not vary from model to model (it does!) but the fact that they are interpreted *outside* the system of models.[24]

Facing the fact that the universal quantifier's interpretation changes from model to model, Sher argues that the interpretation is done outside, that is, "given by *rules external to the system*."[25] I find this move somewhat puzzling, but at the same time encouraging. Why do we think the constancy of a term is guaranteed by the fact that its interpretation is given by external rules? Isn't every interpretation, not just the interpretations of logical constants, given outside the system? I would like to think what Sher meant by "outside the system of models" is the meta-level interpretation, not the first-level interpretation of the universal quantifier, according to my terminology. On the other hand, if we focus on the other level of the meaning of ∀ (say, the meta-level interpretation), that is, 'every,' then Sher's complicated argument would seem quite unnecessary. The meaning constancy of ∀ would have nothing to do with the truth-values of quantified sentences among isomorphic models, and the interpretation 'every' would apply to every model, regardless of the size of the model.

Sher's explanation of ∀ in terms of the Homomorphism Theorem is not well-connected with the rationale for connectives, and I agree with Warmbrōd that it is somewhat awkward and non-motivating to invoke a quite involved criterion for ∀ which is not needed for connectives. On the other hand, I am sympathetic with Sher's project in that it does not ignore a puzzle around ∀: If a domain is the interpretation of a quantifier, it cannot be constant, but ∀ seems to be closer to connectives than to predicates or function symbols in a logical theory.

Sher's project, I believe, is a good example to illustrate how philosophers have tried to grapple with the two different ways that ∀ is taken in its semantic clause. The first level's denotation is a set, and the second level's meta-meaning is 'every.' Sher seems to bridge these two realms and at the same time to find a formal feature of the meta-meaning 'every.' When two models are isomorphic – hence, each domain is of the same size – the difference in the elements of the two domains does not matter in terms of the semantic value of a quantified sentence. Then the meaning of a universal quantifier is more than just its denotation. Instead, there is something constant at a meta-level which guarantees the same semantic value among different models. Sher and some commentators identify this constant feature as the size of

[24]Sher [9, pp. 48–49].

[25]Sher [9, p. 49].

a model. However, the Löwenheim-Skolem Theorem shows us that different sized models could share the constant feature as well, that is, the same semantic value of a first-order sentence among non-isomorphic models. I do not think the same semantic value in different models is due to the role of logical constants, but rather it has something to do with the limit of the expressiveness of first-order logic. Therefore, Sher's proposal, an ambitious and earnest attempt to tackle one of the time-honored controversies as formally as possible, is short of convincing us why quantifiers are logical constants. In the serious defense of the Tarskian analysis, we witness a tension between the two-tiered approaches of the universal quantifier.

Let me visit another place where the tension has become apparent by presenting a fundamental difference between the constancy of connectives and the constancy of the universal quantifier. When we modify classical interpretations of logical constants, and thereby create a non-classical logic, we see a different pattern between connectives and quantifiers in terms of their semantics and inference rules.

In the case of connectives, we pointed out the relation between their assigned meanings and classical/non-classical logic. For example, when we change the meaning of a conditional symbol, a non-classical logic, relevance logic, is born. The semantic clause for $\rightarrow$ is different and a change in inference rules involving $\rightarrow$ occurs. Let us examine what happens if we change the meaning of a quantifier.

When we allow the empty domain as an interpretation of a universal quantifier, we are departing from classical logic to free logic or inclusive logic. According to our account of the two-tiered interpretations of $\forall$, the inclusion of the empty domain means a change in the first-level interpretation. The meta-level meaning 'every' does not seem to change. Note that "free logic is formal logic whose quantifiers are interpreted in the usual way."[26] Nonetheless we need new inference rules in this new inclusive logic.

It is important to make a contrast between relevance logic and free/inclusive logic in terms of what changes and what does not. In the case of relevance logic, a change in the meaning of the conditional symbol (which is expressed in the semantic clause of $(\alpha \rightarrow \beta)$) yields a change in inference rules.[27] Accordingly, we concluded that the meanings of logical constants are constant in a given logic, but could change from logic to logic. That is, their meanings are logic-relative. At the same time we noted that the role of connectives as logical constants is constant from logic to logic, and we therefore do not need to change the definitions of logical concepts – logical consequence, logical truth, and logical equivalence. In the case of free/inclusive logic, only one level (i.e. the first level) of change in terms of the universal quantifier takes place, and its inference rules change. Facing a comparison between connectives and quantifiers in terms of non-classical logics, we have choices to make: (i) We identify the meaning of the universal quantifier at the first level, that is, the domain of discourse. When we include the non-empty domain, a non-classical logic is born. (ii) We take the meta-level interpretation 'every' as

[26] Holt [4, §1.1].

[27] Some might say that a change in inference rules yields a change in meaning.

the meaning of the universal quantifier. Since free/inclusive logic interprets the universal quantifier as 'every' as well, this non-classical logic has nothing to do with a change in the meanings of logical constants.

I claim both views have their own dilemma or more stories to fill in. Position (i) admits that quantifiers are not logical constants since the first-level interpretation is not fixed at all due to different domains of discourse. Considering how much it is accepted that quantifiers are logical constants, I do not think this position would be welcomed by many of us. Position (ii) needs to tell us more about why inference rules involving quantifiers need to change when we do not change the meaning of quantifiers. Especially from a proof-theoretic point of view, it is an inference rule that assigns meanings to logical vocabulary. Then, in this case, the inference rules involving the universal quantifier change but its meaning stays the same. Interestingly enough, both positions end up concluding that free/inclusive logic is not a case where a change in logical constants is the origin of their departure from classical logic. If so, free/inclusive logic stands on a different footing from relevance or intuitionist logic even though both are non-classical logics.

4.4 Concluding Remarks

It is not my view that we should make up our mind about the logical constancy of quantifiers, but that the decision would be a necessary component when we accept the Tarskian standard model-theoretic approach of logical consequence. When we take the logical constant talk seriously (as many do), things become trickier when it comes to quantifiers. To our surprise, we could not find much in common between sentential connectives (which are taken to be logical constants without any controversy) and quantifiers. Then do we need a different kind of rationale for quantifiers? If so, what is it? Our natural language argument intuitively shows the meanings of quantifiers need to be fixed in order to get the right extension of logical consequence. When the formal semantic clause cashes in this intuition, we realize that two different kinds of meaning are assigned to the universal quantifier – a set of objects at the first level and 'every' at a meta-level. A more sophisticated argument for the logical constancy of quantifiers (i.e. the domain size argument) can be viewed as an effort to resolve a tension between these two different levels of the interpretation of quantifiers. The tension is spotted in the case of a non-classical logic in which the empty domain is included as the domain of discourse: The meta-meaning 'every' remains the same and so does the semantic clause of the universal quantified formulas, while the inference rules involving the universal quantifier change.

Suppose I am right that there has been a double way to handle quantifiers. Am I claiming that it is wrong to do so? No, the paper is not about the normative semantics for quantifiers, but aims to be descriptive. I do not see anything wrong with the semantic clause for quantified formulas. Nevertheless the ambiguous feature of quantifier semantics has been overlooked in literature, and I argue that this ambiguity has led to more controversies around quantifiers than around connectives in terms of logical constancy. Does this show that the logical theory (that is, standard model-theory) which assumes the right extension of logical constants is wrong? Not exactly, but the ambiguous nature of quantifiers should give us pause about the entire project whose success hinges on the successful account of logical constants. It would be interesting to inquire further into whether the double feature of quantifiers is reflected in the representational approach. In the case of representational semantics, the meanings are fixed and worlds are changing. It is easy to fix the meaning of 'snow' or 'white' (as we know) and to vary worlds where snow is white or snow is not white. How about 'every' or 'everything'? What does it mean to fix the meaning of 'every'? Does it mean that the domain is whatever we have in the actual world? I would leave the matter here for further work, and just say that the double nature of quantifiers has helped us to look into the nature of logical constants and to get to the bottom of opposing logical views in a fresh way.

References

1. Enderton, H. 2001. *A mathematical introduction to logic*. San Diego/London/Burlington: Harcourt/Academic Press.
2. Etchemendy, J. 1990. *The concept of logical consequence*. Cambridge: Harvard University Press.
3. Gomez-Torrente, M. 1996. Tarski on logical consequence. *Notre Dame Journal of Formal Logic* 37: 125–151.
4. Holt, J. 2010. *Free logic*. Stanford encyclopedia of philosophy. http://plato.stanford.edu/entries/logic-free/
5. MacFarlane, J. 2009. *Logical constants*. Stanford encyclopedia of philosophy. http://plato.stanford.edu/entries/logical-constants/
6. Mostowski, A. 1957. On a generalization of quantifiers. *Fundamenta Mathematicae* 44: 12–35.
7. Peacocke, C. 1976. What is a logical constant? *Journal of Philosophy* 73: 221–240.
8. Shapiro, S. 1998. Logical consequence: Models and modality. In *The philosophy of mathematics today*, ed. M. Schirn, 131–156. Oxford: Clarendon Press.
9. Sher, G. 1991. *The bounds of logic*. Cambridge: MIT.
10. Tarski, A. 1983. *Logic, semantics, metamathematics*, ed. J. Corcoran. Indianapolis: Hackett.
11. Warmbrōd, K. 1999. Logical constants. *Mind* 108: 503–538.

Part II
Semantics of Natural Language

Chapter 5
Conjunctive, Disjunctive, Negative Objects and Generalized Quantification

Ken Akiba

Abstract This paper presents *shadow theory*, according to which, for every object of some type, σ – object in the broadest sense of the term, including truth values and functions – there is another object of the same type, *the negative shadow* of the object, and for every (finite or infinite) set of objects of a single type, σ, there are two other objects of the same type, *the conjunctive shadow* and *the disjunctive shadow* of the set. For instance, Adam has his negative shadow, not-Adam, and Adam and Betty have their conjunctive and disjunctive shadows, Adam-and-Betty and Adam-or-Betty. These are negative, conjunctive, and disjunctive objects in the sense in which they distribute over the objects of one type higher, $\sigma \rightarrow t$, where t is the type of truth values; thus, for instance, Adam-or-Betty is a professor if and only if Adam is a professor or Betty is a professor. The shadows of the same type are divided into infinitely many ranks. The usual infinite hierarchy of types, σ, $\sigma \rightarrow t$, $(\sigma \rightarrow t) \rightarrow t$, $((\sigma \rightarrow t) \rightarrow t) \rightarrow t$, $(((\sigma \rightarrow t) \rightarrow t) \rightarrow t) \rightarrow t$, $\ldots$, are reducible to the bottom two types, σ and $\sigma \rightarrow t$, with the rank distinction among shadows; thus, shadows help simplify type theory. This paper also presents a deductive system based on shadow theory that can be used for the formalization of natural language inferences that involve compound noun phrases and quantifiers. Unlike in Montague's theory or generalized quantifier theory, in this theory such expressions denote objects (shadows) of type e, the type of individuals. There is no essential difference between quantification over individuals (or objects in general) and denotations to them.

K. Akiba (✉)
Department of Philosophy, Virginia Commonwealth University, 915 West Franklin Street, Richmond, VA 23284-2025, USA
e-mail: kakiba@vcu.edu

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_5

5.1 Shadow Theory: Outline

In “General Semantics,” David Lewis wrote thus:

> In the dark ages of logic, a story something like this was told. The phrase ‘some pig’ names a strange thing we may call the *existentially generic pig* which has just those properties that some pig has. Since some pig is male, some pig (a different one) is female, some pig is pink (all over), and some pig is grey (all over), the existentially generic pig is simultaneously male, female, pink, and grey. . . . The phrase ‘every pig’ names a different strange thing called the *universally generic pig* which has just those properties that every pig has. Since not every pig is pink, grey, or any other color, the universally generic pig is not of any color. (Yet neither is he colorless, since not every – indeed not any – pig is colorless.) Nor is he (?) male or female (or neuter), since not every pig is any one of these. He is, however, a pig and an animal, and he grunts; for every pig is a pig and an animal, and grunts. There are also the *negative universally generic pig* which has just those properties that no pig has (he is not a pig, but he is both a stone and a number), the *majority generic pig* which has just those properties that more than half of all pigs have, and many more. . . .
>
> This story is preposterous since nothing, however recondite, can possibly have more or less than one of a set of incompatible and jointly exhaustive properties. At least, nothing can have more or less than one of them *as its properties*. But something, a set, can have *any* combination of them *as its members*; there is no contradiction in that. (Lewis [6], p. 52)

This paper will set forth and defend exactly the kind of theory Lewis called “preposterous.” Josh Gert named it *shadow theory*. According to shadow theory, every individual is paired with another individual, the *negative shadow* of the individual, and every (finite or infinite) set of individuals is paired with two other individuals, its *conjunctive shadow* and *disjunctive shadow*. The pig AyeAye (or A for short) has not-AyeAye (or $\sim A$), and there are such individuals as AyeAye-and-BooBoo ($A \wedge B$), AyeAye-or-BooBoo ($A \vee B$), every-pig ($\forall P$), some-pig ($\exists P$), not-every-pig ($\sim\forall P$), and even no-pig ($\sim\exists P$). These are individuals, and not some higher-order objects such as sets of properties (Lewis) or functions from properties to truth-values (Montague). In the case of propositions, we expect every proposition to have another proposition, its negation, and every (at least finite) set of propositions to have two other propositions, its conjunction and disjunction. Shadow theory generalizes this idea to individuals (and, as you will see, to objects of all types).

The negative shadow of an individual has all and only properties the individual does not have, the conjunctive shadow of a set of individuals has all and only properties all the members of the set have, and the disjunctive shadow of a set of individuals have all and only properties at least some members of the set have. Thus, not-AyeAye is not a pig, is slim, and does not grunt, since AyeAye is a pig, is not slim, and grunts. Some-pig, as Lewis correctly states, is male, is female, is pink all over, and is grey all over. Contrary to Lewis, however, some-pig is not both male and female, nor is it both pink all over and grey all over, since no single pig is both male and female, or is pink all over and grey all over at the same time. Similarly, as Lewis states, it is not the case that every-pig is male, is female, is pink all over, or grey all over. It is not the case, however, that every-pig is neither male nor female, or is not of any color, since every single pig is either male or female, and is of some color. These are immediate consequences of the basic characteristics of shadows just outlined above. What they show is not that shadows are incoherent, but that we have to be a little careful when we reason about shadows; we cannot simply use the inference rules we use for non-shadows for shadows as well. In particular, just because some (disjunctive) shadow has two properties P and Q, we cannot conclude that it also must have the conjunctive property P-and-Q, and just because some (conjunctive) shadow does not have either of two properties P and Q, we cannot conclude that it does not have the disjunctive property P-or-Q. This ought to be the case because some disjunctive shadow may have the properties P and not-P, but it cannot have the conjunctive property P-and-not-P, as neither disjunct has this property (nothing does), and also because some conjunctive shadow may have neither of the properties P and not-P, but it must have the disjunctive property P-or-not-P, as both conjuncts have this property (everything does).

The identities of shadows and other objects, at least for our purposes, are determined classically. That is, on the one hand, classically equivalent objects, such as not-not-AyeAye ($\sim\sim A$) and AyeAye, neither-AyeAye-nor-BooBoo ($\sim(A \vee B)$) and not-AyeAye-and-not-BooBoo ($\sim A \wedge \sim B$), and not-both-AyeAye-and-BooBoo ($\sim(A \wedge B)$) and not-AyeAye-or-not-BooBoo ($\sim A \vee \sim B$), are identical objects; on the other hand, classically non-equivalent objects, such as $\sim A$ and A, and $A \wedge B$ and $A \vee B$, are distinct objects. Thus, the original individuals and their shadows constitute the complete Boolean lattice (or 'algebra'), $B_e = \langle D_e, \wedge, \vee, \sim, 0, 1\rangle$, where D_e is the domain of all individuals, i.e., the original individuals and their shadows, $0 = A \wedge \sim A$, and $1 = A \vee \sim A$ for arbitrary A. (A *complete* Boolean lattice, generally, is a Boolean lattice in which any set – not only finite but also infinite set – of objects has a greatest lower bound, glb or infimum, and a least upper bound, lub or supremum.)

We have written about shadow theory once (see Akiba [1]). There are some differences between the theory presented there and the theory to be presented here, but one major difference concerns this point. The old theory did not consider classically equivalent shadows identical. But that, in hindsight, was a mistake. Generally, the present theory supersedes the old theory.

In this paper, the basic shadow theory outlined above is extended significantly, in two directions. To understand these extensions, we need to have basic understanding of simple type theory, which, thus, will be sketched here. According to simple type theory, every 'object' (in the broadest sense of the term) belongs to one and only one type. There are two kinds of types – basic (or atomic) types and functional types. Examples of basic types are e (the type of individuals), t (the type of truth values), and w (the type of possible worlds). We shall treat expressions and properties purely extensionally in this paper. So we shall use only e and t as the basic types. For every two types (not necessarily distinct) σ and τ, there is a functional type, $\sigma \to \tau$, i.e., the type of functions from the objects of type σ to the objects of type τ. Examples of functional types are $e \to t$ (the type of first-order one-place predicates), $e \to (e \to t)$ (the type of first-order two-place predicates), $e \to (e \to (e \to t))$ (the type of first-order three-place predicates), and so on. (Type notations differ in different traditions and disciplines. The type of individuals is also designated as ι, *i*, or *ind*.

The type of truth values is also designated as o, p, tv, or *bool*. Where we write $\sigma \to (\tau \to \upsilon)$, Church and Montague wrote $(\upsilon\tau)\sigma$ and $\langle\sigma, \langle\tau, \upsilon\rangle\rangle$, respectively.)

Two auxiliary comments may be added to this brief description of simple type theory. First, as is usually the case in the literature, we shall often ignore in this paper the distinction between *sets* and their *characteristic functions*. In the philosophical tradition, a property of individuals, for instance, is extensionally identified with the set that contains those individuals and nothing else. For any set S, the function of type $e \to t$ that gives out the truth value True (or 1) for all the members of S as arguments, and that gives out the truth value False (or 0) for everything else, is called the *characteristic function* of the set S. In mathematics, a set and its characteristic function are often not distinguished because they carry exactly the same information. We shall follow this practice and will not often distinguish one from the other. Thus, the objects of type $e \to t$ may be considered the functions from individuals to truth values, but they may also be considered the sets of individuals. Then, analogously, the objects of type $(e \to t) \to t$ may be considered the sets of sets of individuals, and the objects of type $((e \to t) \to t) \to t$ may be considered the sets of sets of sets of individuals, and so on.

Second, as you can see above, *many-place* predicates are treated differently in simple type theory than in the ordinary logic; for instance, first-order two-place predicates are treated as denoting second-order functions from individuals to first-order one-place predicates, and, generally, mth-order n-place predicates are treated as denoting $(m + 1)$th-order functions from $(m - 1)$th-order objects to mth-order $(n - 1)$-place predicates. This conception is plausible because from an mth-order n-place predicate we create an mth-order $(n - 1)$-place predicate by filling in one of its variables with an $(m - 1)$th-order object. In this conception, mth-order n-place predicates are in fact of a type higher than that of mth-order $(n - 1)$-place predicates. We can dispense with many-place predicates in formal grammar as primitive entities in favor of one-place predicates and higher-order functions. This idea is embraced also in this paper; our formal language contains only one-place predicates and no many-place predicates.

We are now in a position to understand the two-way extension of the basic shadow theory, to be proposed in this paper. First, we have just seen that every individual (i.e., object of type e) has its negative shadow, and that every set of individuals has its conjunctive and disjunctive shadows. In the extended version, however, shadows are not confined to those of individuals. In fact, *anything* of *any* type – not just every *individual* but every *truth value* and every *function* as well – has its negative shadow of the same type, and *any* (finite or infinite) set of things of one and the same type, *any* type, is paired with its conjunctive and disjunctive shadows of the same type. The original objects of type σ and their shadows constitute a complete Boolean algebra $B_\sigma = \langle D_\sigma, \wedge_\sigma, \vee_\sigma, \sim_\sigma, 0_\sigma, 1_\sigma\rangle$ just as the original individuals and their shadows do. (The complete Boolean algebra of type e may now be designated as $B_e = \langle D_e, \wedge_e, \vee_e, \sim_e, 0_e, 1_e\rangle$.) Just as a shadow of individuals, of type e, distributes its constituents over first-order properties, of type $e \to t$ (e.g., AyeAye-and-BooBoo grunt(s) if and only if AyeAye grunts and BooBoo grunts), a shadow of objects of type σ distributes its constituents over properties of the higher type $\sigma \to t$. For instance, a shadow of first-order properties, of type $e \to t$, distributes its constituents over second-order properties, of type $(e \to t) \to t$. Thus, the conjunctive shadow $P \wedge Q$ of first-order properties P and Q is not the same as the intersection $P \cap Q$ of the properties (understood extensionally), for the intersection distributes its constituents, P and Q, over the objects of the *lower* type, e. For instance, being a grey female, $G \wedge F$, is an attractive feature if and only if being grey is an attractive feature and being a female is an attractive feature (assuming that attractiveness distributes), whereas the individual AyeAye is a grey female, i.e., AyeAye is in $G \cap F$, if and only if AyeAye is grey and AyeAye is a female. This distinction between negative, conjunctive, and disjunctive shadows, $\sim P$, $P \wedge Q$, and $P \vee Q$, on the one hand, and complements, intersections, and unions, $\overline{P}$, $P \cap Q$, and $P \cup Q$, on the other, ought to be carefully drawn and understood.

In the second extension, each type is divided into infinitely many *ranks*, and any object of any type at any rank has not only its negative shadow of the same type at the same rank, but also its negative shadow of the same type at any higher rank, and any set of objects of any type at any rank has not only its conjunctive and disjunctive shadows of the same type at the same rank, but also its conjunctive and disjunctive shadows of the same type at any higher rank; and those shadows at higher ranks are all distinct from one another and distinct from their lower-ranked counterparts. The basic idea can be understood intuitively as follows. Take the type e of individuals for instance; designate their original domain D_e^0. Again, they and their shadows constitute a complete Boolean algebra. Let's designate it as $B_e^2 = \langle D_e^2, \wedge_e^2, \vee_e^2, \sim_e^2, 0_e^2, 1_e^2 \rangle$, determined by the partial ordering $\leq_e^2$. (For reasons that will become clear shortly, we do not use 1 for the superscript here.) But, now, take all the objects in D_e^2 (some of which are original individuals also in D_e^0), forget the partial ordering $\leq_e^2$, and suppose they have their own shadows and constitute another complete Boolean algebra, $B_e^4 = \langle D_e^4, \wedge_e^4, \vee_e^4, \sim_e^4, 0_e^4, 1_e^4 \rangle$, determined by its own partial ordering $\leq_e^4$. Then, take all the objects in D_e^4, forget the partial ordering $\leq_e^4$, and suppose they have their own shadows and constitute yet another complete Boolean algebra $B_e^6 = \langle D_e^6, \wedge_e^6, \vee_e^6, \sim_e^6, 0_e^6, 1_e^6 \rangle$, and so on and on, indefinitely. The result is what we call an *infinitely nested (complete) Boolean structure* or *the infinite (complete) Boolean expansion of* D_e^0. Furthermore, we suppose that analogous nested structures exist in all types. Thus, for instance, $\sim_e^4 \sim_e^2 A_e^0$ is a shadow different from $\sim_e^2 \sim_e^2 A_e^0 = \sim_e^4 \sim_e^4 A_e^0 = A_e^0$, and there are no such shadows as $\sim_e^2 \sim_e^4 A_e^0$ and $\sim_e^2 \sim_e^2 A_e^4$; $\sim_{e\to t}^5 (A_{e\to t}^1 \vee_{e\to t}^5 B_{e\to t}^3)$ is the same shadow as $\sim_{e\to t}^5 A_{e\to t}^1 \wedge_{e\to t}^5 \sim_{e\to t}^5 B_{e\to t}^3$, but not as $\sim_{e\to t}^3 A_{e\to t}^1 \wedge_{e\to t}^3 \sim_{e\to t}^3 B_{e\to t}^3$.

These are the two extensions of the basic shadow theory, to be fleshed out in the rest of this paper. Admittedly, the ontology of shadows resulting from the two-way extension is huge. However, the following theorem holds:

Type Reduction Theorem. *For any type σ and rank n that exists in the type, the objects of type $(\sigma \to t) \to t$ at rank n, where t (again) is the type of truth values, are identifiable and reducible to the objects of type σ at rank n. Furthermore, by repeating the same reduction procedure, we can reduce types $(\sigma \to t) \to t$, $((\sigma \to t) \to t) \to t$, $(((\sigma \to t) \to t) \to t) \to t$, $((((\sigma \to t) \to t) \to t) \to t) \to t$, etc., to*

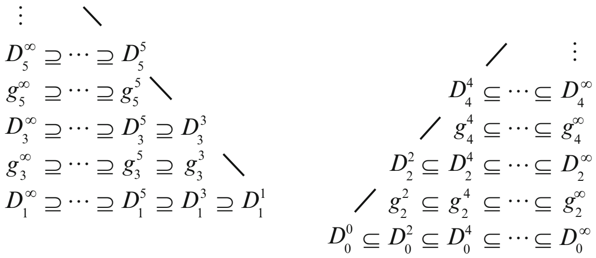


Fig. 5.1 Type reduction

two types, σ and $\sigma \to t$. Thus, we do not really need those higher-order types; we need only the two types, σ and $\sigma \to t$, with infinitely many ranks.

This reduction is graphically expressed in Fig. 5.1 above, where the number n in the subscripts is an abbreviation of $(\cdots((\underbrace{\sigma \to t) \to t) \to \cdots) \to t)}_{n \text{ times}} \cdots)$ – thus, $D_0^0 = D_\sigma^0$, $D_1^1 = D_{\sigma\to t}^1$, $D_2^2 = D_{(\sigma\to t)\to t}^2$, etc. – and each g_n^m is a one-to-one correspondence between the domain D_n^m above it and the domain D_{n-2}^m below it. (You may now see the reason why we stipulated that the superscript for each domain be increased by two, not one, from that for the original domain. The meaning of '$g_n^m \subseteq g_n^{m+2}$' will be revealed later.) What the Type Reduction Theorem states is that all the types in the figure except the two at the bottom are not necessary. So, by expanding shadows at the bottom, we can eliminate the infinite hierarchy of types. The initially large ontology of shadows can be reduced significantly, along with the infinite hierarchy of types that must otherwise exist. This result indicates the great potential of shadow theory to simplify ontology.

This completes the outline of shadow theory, to be spelled out fully in the rest of this paper. In Sect. 5.2, Montague's higher-order treatment of noun phrases is examined, and its problems are pointed out, which will give motivations for shadow theory. In Sect. 5.3, shadow theory is formalized, and the Type Reduction Theorem is proved. In Sect. 5.4, a sequent calculus based on shadow theory is presented, and some natural language examples are given. Section 5.5 summarizes the discussion.

5.2 Background: Montague's Treatment of Noun Phrases

One major question that concerns this paper is how to formalize sentences like the following:

1. Adam is a professor.
2. Adam and Betty are professors.
3. Adam or Betty is a professor.
4. Not Adam but Betty is a professor.

5. Not both Adam and Betty are professors.
6. Neither Adam nor Betty is a professor.
7. Everybody in the audience is a professor.
8. Somebody in the audience is a professor.
9. Nobody in the audience is a professor.

The founders of modern logic, Gottlob Frege and Bertrand Russell, warned us against uniform treatment of all noun phrases (NPs), phrases such as *Adam, Adam and Betty, Adam or Betty, everybody in the audience,* and *somebody in the audience* above. They thought it a mistake to consider *Adam is a professor, Adam and Betty are professors, Adam or Betty is a professor, Everybody in the audience is a professor*, and *Somebody in the audience is a professor* to have the same subject-predicate form; they thought it a mistake to think that *Adam and Betty, Adam or Betty, everybody in the audience*, and *somebody in the audience* denote individuals just as *Adam* does. In their formalization, (1), (2), (3), (7), and (8) would look differently as follows:

10. Pa.
11. $Pa \wedge Pb$.
12. $Pa \vee Pb$.
13. $\forall x(Ax \rightarrow Px)$.
14. $\exists x(Ax \wedge Px)$.

(Needless to say, their original notations were different.) There is no grammatical unit in (11) through (14) that corresponds to *Adam and Betty, Adam or Betty, everybody in the audience*, and *somebody in the audience*. As is well known, Russell [11] did not assign any distinct semantic value even to such definite descriptions as *the Queen of England* and *the President of the United States.*

The problem with this approach, however, is that it ignores the compositional structures of sentences in natural language. This is particularly problematic in the eyes of formal semanticists of natural language, who generally embrace the *compositionality principle*: the semantic value (or meaning) of a compound expression is to be determined by the semantic values of its constituents in accordance with its grammatical structure. Surely, the meaning of *Everybody (Somebody) in the audience is a professor* does seem to be determined by the meaning of *everybody (somebody) in the audience* and the meaning of *is a professor* plus the subject-predicate structure of the sentence, just as the meaning of *Adam is a professor* is determined by the meaning of *Adam* and the meaning of *is a professor* plus the subject-predicate structure of the sentence.

One of Richard Montague's greatest contributions to formal semantics, facilitated by Church's [4] development of λ-calculus, is his discovery of a way to assign semantic values uniformly to all NPs. According to Montague [7], *Adam and Betty* is given as its semantic value the function from the properties both Adam and Betty possess to the truth value True, and from the other properties to the truth value False; *Adam or Betty* is given as its semantic value the function from the properties either Adam or Betty possesses to the truth value True, and from the other properties to the truth value False; *everybody in the audience* is given as its semantic value the function from the properties everybody in the audience possesses to the truth value True, and from the other properties to the truth value False; *somebody in the audience* is given as its semantic value the function from the properties at least one person in the audience possesses to the truth value True, and from the other properties to the truth value False; and so on. These are functions of type $(e \to t) \to t$ because a property of individuals, such as that of being a professor, is considered a function from each individual to the truth value True or False, depending on whether the individual possesses that property, and, thus, is of type $e \to t$.

More formally,

15. $[\![\textit{Adam and Betty}]\!] = [\![\lambda X.Xa \wedge Xb]\!]$;
16. $[\![\textit{Adam or Betty}]\!] = [\![\lambda X.Xa \vee Xb]\!]$;
17. $[\![\textit{everybody in the audience}]\!] = [\![\lambda X.\forall x(Ax \to Xx)]\!]$;
18. $[\![\textit{somebody in the audience}]\!] = [\![\lambda X.\exists x(Ax \wedge Xx)]\!]$,

whereas the semantic value of *is a professor* (or *are professors*) is the property of being a professor:

19. $[\![\textit{is a professor}]\!] = [\![\lambda x.Px]\!]$.

(The semantic value of the expression x will be indicated as $[\![x]\!]$ in this paper. We do not make a distinction between *being a professor* and *being professors*. We shall treat properties purely extensionally for the sake of simplicity.) Then *Adam and Betty are professors, Adam or Betty is a professor*, and *Everybody (Somebody) in the audience is a professor* are true if and only if the semantic values of the above NPs, i.e., functions from properties to truth values, give True to the argument, the property of being a professor:

20. $[\![\textit{Adam and Betty is a professor}]\!] = \mathrm{T}$
    iff $[\![\textit{Adam and Betty}]\!]([\![\textit{is a professor}]\!]) = \mathrm{T}$
    iff $[\![(\lambda X.Xa \wedge Xb)(\lambda x.Px)]\!] = [\![(\lambda x.Px)a \wedge (\lambda x.Px)b]\!] = [\![Pa \wedge Pb]\!] = \mathrm{T}$
21. $[\![\textit{Adam or Betty is a professor}]\!] = \mathrm{T}$
    iff $[\![\textit{Adam or Betty}]\!]([\![\textit{is a professor}]\!]) = \mathrm{T}$
    iff $[\![(\lambda X.Xa \vee Xb)(\lambda x.Px)]\!] = [\![(\lambda x.Px)a \vee (\lambda x.Px)b]\!] = [\![Pa \vee Pb]\!] = \mathrm{T}$
22. $[\![\textit{Everybody in the audience is a professor}]\!] = \mathrm{T}$
    iff $[\![\textit{everybody in the audience}]\!]([\![\textit{is a professor}]\!]) = \mathrm{T}$
    iff $[\![(\lambda X.\forall x(Ax \to Xx))(\lambda x.Px)]\!] = [\![\forall x(Ax \to (\lambda x.Px)x)]\!] = [\![\forall x(Ax \to Px)]\!] = \mathrm{T}$
23. $[\![\textit{Somebody in the audience is a professor}]\!] = \mathrm{T}$
    iff $[\![\textit{somebody in the audience}]\!]([\![\textit{is a professor}]\!]) = \mathrm{T}$
    iff $[\![(\lambda X.\exists x(Ax \wedge Xx))(\lambda x.Px)]\!] = [\![\exists x(Ax \wedge (\lambda x.Px)x)]\!] = [\![\exists x(Ax \wedge Px)]\!] = \mathrm{T}$.

This shows that Montague's interpretations give the same truth conditions to the relevant sentences as the Frege-Russell interpretations do. The idea sketched above helped Montague to construct semantics for a fragment of English that assigns a semantic value to each grammatical unit in accordance with its compositional structure.

Furthermore, Montague treats *all* NPs uniformly as expressions of type $(e \to t) \to t$. The type $(e \to t) \to t$ is the type of second-order one-place predicates (or second-order properties), but it is also often called 'the type of (first-order) quantifiers' because quantifiers such as *everybody in the audience, somebody in the audience*, and *nobody in the audience* are all of this type. In Montague's approach, followed by *generalized quantifier theory* (see, e.g., Barwise and Cooper [3]), all NPs (at least in subject position) are considered quantifiers of sort. An important consequence of Montague's uniform treatment of NPs is that even proper names such as *Adam* and *Betty* are considered to be of the same type and denote not individuals but functions from first-order properties to truth-values. For instance, *Adam* denotes the function from the set of property Adam possesses to the truth value True and from the other properties to the truth value False. Thus,

24. $[\![Adam]\!] = [\![\lambda X.Xa]\!]$.

Adam is a professor is analyzed accordingly:

25. $[\![Adam]\!]([\![is\ a\ professor]\!]) = \mathrm{T}$
    iff $[\![(\lambda X.Xa)(\lambda x.Px)]\!] = [\![(\lambda x.Px)a]\!] = [\![Pa]\!] = \mathrm{T}$.

Compare this with the Frege-Russell interpretation of the same sentence:

26. $[\![is\ a\ professor]\!]([\![Adam]\!]) = \mathrm{T}$
    iff $[\![(\lambda x.Px)a]\!] = [\![Pa]\!] = \mathrm{T}$.

While the truth conditions of the sentence, as well as the semantic values of the predicate *is a professor*, are the same on Montague's and Frege and Russell's analyses, the semantic values of *Adam* are different, and the way the truth conditions are determined is quite the reverse. For Frege and Russell, *Adam* is of type e and denotes the individual Adam, and *is a professor* is of the higher type $e \to t$ and takes *Adam* as an argument of the function and gives out the truth value True, of type t, whereas for Montague, Adam is of type $(e \to t) \to t$ and denotes a second-order function that takes *is a professor* of type $e \to t$ as its argument and gives out the truth value True.

In the Montague-generalized quantifier tradition, single names such as *Adam* and *Betty*, as well as compound NPs such as *Adam and Betty* and *Adam or Betty*, are all treated as generalized quantifiers, i.e., treated in the same way as quantifiers in the narrower sense, such as *everybody in the audience* and *somebody in the audience*. Hence the title of this paper.

On Montague's analysis, no proper name denotes an individual. Indeed, most likely nothing does (though how Montague will treat demonstratives and pronouns such as *this, that, he,* and *she* is somewhat unclear). This is quite counterintuitive and unpalatable. It is clearly in opposition to Kripke's [5] influential view that proper names are rigid designators that denote individuals directly without any mediation of their properties. But even setting aside Kripke's view, it is really strange that we cannot denote any individual in English. This difficulty with Montague's analysis is often obscured but also compounded by the fact that in the metalanguage, the language of semantics, *Adam* does denote the individual Adam, and not the set of properties Adam possesses. For instance, when we say that the English proper name *Adam* denotes the function from the set of property Adam possesses to the truth value True and from everything else to the truth value False (or $[\![\lambda X.Xa]\!]$), the second, metalinguistic token of 'Adam' (or a) is supposed to denote the individual Adam, and not the function in question, for, otherwise, the account will be circular and we cannot anchor the denotation in the world; put simply, we cannot determine what individual we are talking about. But if we can denote the individual Adam by 'Adam' in the language of semantics (which itself may well be English), why should we not be able to do so in the object language, English?

Barbara Partee ([9]; also see Partee and Rooth [10]) embraced the general Montagovian approach to semantics but dissented from Montague's uniform treatment of NPs. According to Partee, NPs belong to at least three different types, e, $e \to t$, and $(e \to t) \to t$, depending on the context. Compare:

27. John/the man/a man walked in. He looked tired.
28. Mary considers John competent in semantics and an authority on unicorns.
29. Every/a man is rational.

John/the man/a man in (27) seems to be of type e (what Partee calls "referential NPs"), while *every/a man* in (29) seems to be of type $(e \to t) \to t$ ("quantificational NPs"). *An authority on unicorns* in (28) seems to be of type $e \to t$ ("predicative NPs"), for it is conjoined with the adjective phrase (AP) *competent in semantics*, which, denoting a function from individuals to truth values, is of that type. Note that in these examples, even the same NP *a man* can be of different types. Partee set forth *type-shifting principles*, which allow the same NPs to shift their types in accordance with the context. Thus, according to Partee, *Adam* in *Adam is a professor* is of type e, but *Adam* in *Adam and Betty are professors* is of type $(e \to t) \to t$, formalized respectively thus:

30. $[\![\textit{is a professor}]\!]([\![\textit{Adam}]\!])$;
31. $[\![\textit{Adam and Betty}]\!]([\![\textit{is a professor}]\!])$.

The type of *Adam* is lifted from e in (30) to $(e \to t) \to t$ in (31). This operation is called *type-lifting*.

Partee's treatment of NPs, however, is no more satisfying than Montague's. Whatever differences one might find among the NPs in (27)–(29), there are no similar differences between the NPs in *Adam is a professor* and *Adam and Betty are professors*. It is quite counterintuitive to consider *Adam* in these two sentences to play totally different roles, one the referential role, denoting an individual, an argument of a function, and the other the quantificational role, denoting a function. (Conversely, the predicate *is a professor* in the former sentence must play as a function, and that in the latter as an argument of a function.)

The same problem arises between proper names and pronouns if one takes proper names as quantifiers but pronouns, such as *he* and *this*, to denote individuals, as one might be tempted to do to rescue Montague's approach. For conjoined NPs such as *he and Betty* and *he or Betty* will have to be quantifiers; so the type of the pronoun *he* cannot be the same when it appears singly or as a part of conjoined NPs.

Thus, we have a dilemma. On the one hand, to anchor denotations in the world proper names must belong to the lowest type, type *e*, denoting individuals. On the other hand, conjoinability considerations would raise the type of proper names to the type of quantifiers. Obviously the only way to avoid this dilemma is to take all NPs, including quantifiers, to belong to the type of individuals, and that's what shadow theory proposes to do. (Or, at the very least, that's the guiding principle. In the end there may remain some recalcitrant NPs, e.g., those in the 'donkey' sentences, which may resist this analysis.)

After shadow theory is spelled out in the next section, in Sect. 5.4 a formal language that involves denotations to shadows is presented. In that language, the sentences (1)–(9) above is formalized as follows, where *A* in (32) through (37) is *Adam*, *B* is *Betty*, *P* is *is a professor*, and *A* in (38)–(40) is *is a person in the audience*:

32. $(P^1_{\sim e}A^0_e)^0_\bot$.
33. $(P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\bot$.
34. $(P^1_{\sim e}(A^0_e \vee B^0_e)^2_e)^0_\bot$.
35. $(P^1_{\sim e}((\sim A^0_e)^2_e \wedge B^0_e)^2_e)^0_\bot$.
36. $(P^1_{\sim e}(\sim(A^0_e \wedge B^0_e)^2_e)^2_e)^0_\bot$ or $(P^1_{\sim e}((\sim A^0_e)^2_e \vee (\sim B^0_e)^2_e)^2_e)^0_\bot$.
37. $(P^1_{\sim e}(\sim(A^0_e \vee B^0_e)^2_e)^2_e)^0_\bot$ or $(P^1_{\sim e}((\sim A^0_e)^2_e \wedge (\sim B^0_e)^2_e)^2_e)^0_\bot$.
38. $(P^1_{\sim e}(\forall A^1_{\sim e})^2_e)^0_\bot$.
39. $(P^1_{\sim e}(\exists A^1_{\sim e})^2_e)^0_\bot$.
40. $(P^1_{\sim e}(\sim(\exists A^1_{\sim e})^2_e)^2_e)^0_\bot$.

Here the subscripts indicate types, where $\sim e = e \to t$ and $\bot = t$, and the superscripts indicate ranks. In practice, most of the indices can be omitted. Ignore most of the indices at this point, and focus on the structures and the fact that all the NPs involved are assigned type *e*. As you can see here, in the language based on shadow theory, NPs, including quantifiers, can uniformly denote individuals of type *e*, i.e., the original individuals and their shadows. The quantifiers '$\forall A$', '$\exists A$', and '$\sim(\exists A)$' read *everybody/somebody/nobody in the audience*, and denote the conjunctive shadow/the disjunctive shadow/the negative shadow of the disjunctive shadow of all the people in the audience, respectively.

Incidentally, as is already clear above, the symbols $\wedge$, $\vee$, and $\sim$ are used in various ways in this paper. In particular, they are used in the metalanguage in the descriptions of conjunctive, disjunctive, and negative shadows, and they are also used in the formal, object language, e.g., in sentences like the above, denoting shadows. They are also used to indicate glbs, lubs, and complements in Boolean algebras. (Shadows are special instances of glbs, lubs, and complements.) Sometimes they are also used in the narrative as the ordinary sentential conjunction, disjunction, and negation. There are three main reasons why the symbols are used thus, despite the potential ambiguity. First, the central point of shadow theory is that all kinds of things can be conjoined, disjoined, and negated, not just propositions. What's common among the conjunctions, disjunctions, and negations of different kinds is that they all constitute Boolean algebras. Second, one virtue of shadow theory is that shadows closely correspond to expressions (such as NPs) in natural language, so much so that we do not need much distinction between the metalanguage and the object language. Finally, it would be only cumbersome, with very little to gain, to distinguish all those, slightly different conjunctions, disjunctions, and negations with different symbols. For these reasons, $\wedge$, $\vee$, and $\sim$ are used in the diverse ways in this paper.

5.3 Formalization of Shadow Theory and the Type Reduction Theorem

In this section, we shall consider a formalized account of shadow theory, and prove the Type Reduction Theorem mentioned in Sect. 5.1.

Consider any type σ. For any non-negative integer n, the domain D_n^n of type $(\cdots((\underbrace{\sigma \to t) \to t) \to \cdots) \to t)}_{n \text{ times}}\cdots)$ is defined thus:

- $D_0^0 = D_\sigma \neq \emptyset$.
- $D_t^0 = D_t = \{0, 1\}$.
- $D_n^n = \{f : f \text{ is a function from } D_{n-1}^{n-1} \text{ to } D_t^0\}$ if $n \neq 0$.

For any two even (or odd) integers $m > 1$ and $n \geq 0$, D_n^m is the domain of the complete Boolean algebra $B_n^m = \langle D_n^m, \leq_n^m \rangle = \langle D_n^m, \wedge_n^m, \vee_n^m, \sim_n^m, 0_n^m, 1_n^m \rangle$, where

- $D_n^{m-2} \subseteq D_n^m$;
- $\forall x, y \in D_n^{m-2}. x \not\leq_n^m y \wedge y \not\leq_n^m x$;
- $\forall x \in (D_n^m - D_n^{m-2})$. x is defined by some y's $\in D_n^{m-2}$ and $\wedge_n^m$, $\vee_n^m$, and $\sim_n^m$;
- $\forall x, y \in D_n^m.(x \leq_n^m y \wedge y \leq_n^m x) \leftrightarrow x = y$;
- $\forall x, y \in D_n^m. x \leq_n^m y \leftrightarrow (\forall z \in D_{n+1}^{m-1}. zx \leq_t^0 zy)$; [1]

where 'zx' and 'zy' are simplified expressions of the functions $z(x)$ and $z(y)$. In what follows, similar parentheses are omitted. Also, constants and variables that belong to some domains are expressed with sub- and superscripts; for instance, the last sentence above may be abbreviated as '$\forall x_n^m, y_n^m. x_n^m \leq_n^m y_n^m \leftrightarrow (\forall z_{n+1}^{m-1}. z_{n+1}^{m-1} x_n^m \leq_t^0 z_{n+1}^{m-1} y_n^m)$'. The numbers in the superscripts, m's, are called *ranks*. Because $D_n^{m-2} \subseteq D_n^m$, an object that belongs to some type at some rank m belongs also to the same type at all the ranks higher than m, i.e., $m + 2k$ for any k. The lowest of such ranks is called *the rank of the object*.

The third item above is only informally presented. It is not very easy to formalize it using mathematical induction, and it is not worth the effort. Its point ought to be sufficiently clear: we can treat y's in D_n^{m-2} like atomic statements and make compound statements using $\wedge_n^m$, $\vee_n^m$, and $\sim_n^m$; then those are the only objects added in $D_n^m - D_n^{m-2}$. The fourth item states that those classically equivalent are indeed identical.

Finally,

- $D_n^\infty = \bigcup_{0 \leq m}^{\infty} D_n^{n+2m}$

Again, D_n^∞ may be called an *infinitely nested (complete) Boolean structure* or the *infinite (complete) Boolean expansion of* D_n^n. Obviously, such expansions may be continued to the transfinite levels; however, we shall make no such expansion in this paper because that seems to have little significance in the analysis of natural language, which is our main concern.

Intuitively, what is described above is this: we start with the series of domains $D_0^0, D_1^1, D_2^2, D_3^3, \ldots$, where $D_0^0 \neq \emptyset$ and D_n^n is the set of functions from D_{n-1}^{n-1} to $\{0, 1\}$, and expand each into D_n^{n+2} by adding (only) glbs, lubs, and complements with respect to the partial order $\leq_n^{n+2}$, then further expand it into D_n^{n+4} by adding (only) glbs, lubs, and complements with respect to the partial order $\leq_n^{n+4}$, and so on, indefinitely. The result is the infinitely nested Boolean structure D_n^∞. See, again, Fig. 5.1. As a consequence of [1], for any $x, y \in D_n^m$ and $z \in D_{n+1}^{m-1}$,

$$z_{n+1}^{m-1}(\sim_n^m x_n^m) =_t^0 \sim_t^0 (z_{n+1}^{m-1} x_n^m);$$

$$z_{n+1}^{m-1}(x_n^m \wedge_n^m y_n^m) =_t^0 z_{n+1}^{m-1} x_n^m \wedge_t^0 z_{n+1}^{m-1} y_n^m;$$

$$z_{n+1}^{m-1}(x_n^m \vee_n^m y_n^m) =_t^0 z_{n+1}^{m-1} x_n^m \vee_t^0 z_{n+1}^{m-1} y_n^m;$$

$$z_{n+1}^{m-1}(x_n^m \rightarrow_n^m y_n^m) =_t^0 z_{n+1}^{m-1} x_n^m \rightarrow_t^0 z_{n+1}^{m-1} y_n^m;$$

where $x_n^m \rightarrow_n^m y_n^m =_{df} \sim_n^m x_n^m \vee_n^m y_n^m$ and $x_n^m =_n^m y_n^m =_{df} (x_n^m \rightarrow_n^m y_n^m) \wedge_n^m (y_n^m \rightarrow_n^m x_n^m)$. Put simply, z_{n-1}^{m+1} is a homomorphism from the Boolean algebra B_n^m to the two-element Boolean algebra B_t^0. The negative, conjunctive, and disjunctive objects involved in the left-hand side of the equations may be considered negative, conjunctive, and disjunctive shadows, respectively. The equations assert that the shadows distribute over the objects of one type higher.

Let us now define two functions f_n^m and g_n^m. Define $f_n^m : D_n^m \rightarrow D_{n-1}^{m+1}$ thus:

$$\lambda f_n^m . \forall x_{n-1}^{m-1} . \forall y_n^m . y_n^m x_{n-1}^{m-1} =_t^0 y_n^m (f_n^m y_n^m \rightarrow_{n-1}^{m+1} x_{n-1}^{m-1}). \qquad [2]$$

In particular, if y_n^m is the characteristic function of the set $(Sy)_{n-1}^{m-1} \subseteq D_{n-1}^{m-1}$, then $f_n^m y_n^m = \bigwedge_{n-1}^{m+1} \{x_{n-1}^{m-1} \text{ if } x_{n-1}^{m-1} \in (Sy)_{n-1}^{m-1}; \sim_{n-1}^{m+1} x_{n-1}^{m-1} \text{ if } x_{n-1}^{m-1} \notin (Sy)_{n-1}^{m-1}\}$. For instance, if $D_0^0 = \{a, b, c\}$ and $(Sy)_0^0 = \{a\}$, then $f_1^1 y_1^1 = a_0^0 \wedge_0^2 \sim_0^2 b_0^0 \wedge_0^2 \sim_0^2 c_0^0$.

Then define $g_n^m : D_n^m \rightarrow D_{n-2}^m$ thus:

$$\lambda g_n^m . \forall y_{n-1}^{m-1} . \forall z_n^m . z_n^m y_{n-1}^{m-1} =_t^0 y_{n-1}^{m-1} (f_{n-1}^{m-1} y_{n-1}^{m-1} \rightarrow_{n-2}^m g_n^m z_n^m). \qquad [3]$$

In particular, if z_n^m is the characteristic function of the set $(Sz)_{n-1}^{m-1} \subseteq D_{n-1}^{m-1}$ and if $y_{n-1}^{m-1} \in (Sz)_{n-1}^{m-1}$ is the characteristic function of the set $(Sy)_{n-2}^{m-2} \subseteq D_{n-2}^{m-2}$, then $g_n^m z_n^m = \bigvee_{n-2}^m \bigwedge_{n-2}^m \{x_{n-2}^{m-2} \text{ if } x_{n-2}^{m-2} \in (Sy)_{n-2}^{m-2}; \sim_{n-2}^m x_{n-2}^{m-2} \text{ if } x_{n-2}^{m-2} \notin (Sy)_{n-2}^{m-2}\}$. For instance, ignoring the difference between a set and its characteristic function, if $D_0^0 = \{a, b, c\}$ and $(Sz)_1^1 = \{\{a\}, \{a, b\}, \{b, c\}\}$, then $g_2^2 z_2^2 = (a_0^0 \wedge_0^2 \sim_0^2 b_0^0 \wedge_0^2 \sim_0^2 c_0^0) \vee_0^2 (a_0^0 \wedge_0^2 b_0^0 \wedge_0^2 \sim_0^2 c_0^0) \vee_0^2 (\sim_0^2 a_0^0 \wedge_0^2 b_0^0 \wedge_0^2 c_0^0)$. That is, intuitively, g_n^m may be thought of as the function that constructs a shadow in *disjunctive normal form* from the corresponding set of sets.

We shall prove that g_n^m is a one-to-one correspondence from D_n^m to D_{n-2}^m. For this goal, we prove

Lemma 5.1. *For any* $y_n^m, f_n^m y_n^m =_{n-1}^{m+3} f_n^{m+2} y_n^m$. *(We express this as* $f_n^m \subseteq f_n^{m+2}$.*)*

Proof. Consider $f_n^{m+2} : D_n^{m+2} \to D_n^{m+3}$.

$$\forall x_{n-1}^{m+1}.\forall y_n^{m+2}.y_n^{m+2} x_{n-1}^{m+1} =_t^0 y_n^{m+2}(f_n^{m+2} y_n^{m+2} \to_{n-1}^{m+3} x_{n-1}^{m+1}).$$

Since $D_{n-1}^{m-1} \subseteq D_{n-1}^{m+1}$ and $D_n^m \subseteq D_n^{m+2}$,

$$\forall x_{n-1}^{m-1}.\forall y_n^m.y_n^m x_{n-1}^{m-1} =_t^0 y_n^m(f_n^{m+2} y_n^m \to_{n-1}^{m+3} x_{n-1}^{m-1}).$$

Comparing this with [2], for any x_{n-1}^{m-1} and y_n^m, $y_n^m(f_n^m y_n^m \to_{n-1}^{m+1} x_{n-1}^{m-1}) =_t^0 y_n^m(f_n^{m+2} y_n^m \to_{n-1}^{m+3} x_{n-1}^{m-1})$. So $y_n^m(f_n^m y_n^m) \to_t^0 y_n^m x_{n-1}^{m-1} =_t^0 y_n^m(f_n^{m+2} y_n^m) \to_t^0 y_n^m x_{n-1}^{m-1}$. Thus, $y_n^m(f_n^m y_n^m) =_t^0 y_n^m(f_n^{m+2} y_n^m)$. Therefore, for any $y_n^m, f_n^m y_n^m =_{n-1}^{m+3} f_n^{m+2} y_n^m$. □

Theorem 5.1. *For any* z_n^m, $g_n^m z_n^m =_{n-2}^{m+2} g_n^{m+2} z_n^m$. *(We express this as* $g_n^m \subseteq g_n^{m+2}$.*)*

Proof. Consider $g_n^{m+2} : D_n^{m+2} \to D_{n-2}^{m+2}$.

$$\forall y_{n-1}^{m+1}.\forall z_n^{m+2}.z_n^{m+2} y_{n-1}^{m+1} = y_{n-1}^{m+1}(f_{n-1}^{m+1} y_{n-1}^{m+1} \to_{n-2}^{m+2} g_n^{m+2} z_n^{m+2}).$$

Since $D_{n-1}^{m-1} \subseteq D_{n-1}^{m+1}$ and $D_n^m \subseteq D_n^{m+2}$,

$$\forall y_{n-1}^{m-1}.\forall z_n^m.z_n^m y_{n-1}^{m-1} = y_{n-1}^{m-1}(f_{n-1}^{m+1} y_{n-1}^{m-1} \to_{n-2}^{m+2} g_n^{m+2} z_n^m).$$

By Lemma 5.1,

$$\forall y_{n-1}^{m-1}.\forall z_n^m.z_n^m y_{n-1}^{m-1} = y_{n-1}^{m-1}(f_{n-1}^{m-1} y_{n-1}^{m-1} \to_{n-2}^{m+2} g_n^{m+2} z_n^m).$$

Comparing this with [3], for any y_{n-1}^{m-1} and z_n^m, $y_{n-1}^{m-1}(f_{n-1}^{m-1} y_{n-1}^{m-1} \to_{n-2}^m g_n^m z_n^m) =_t^0 y_{n-1}^{m-1}(f_{n-1}^{m-1} y_{n-1}^{m-1} \to_{n-2}^{m+2} g_n^{m+2} z_n^m)$. So $y_{n-1}^{m-1}(f_{n-1}^{m-1} y_{n-1}^{m-1}) \to_t^0 y_{n-1}^{m-1}(g_n^m z_n^m) =_t^0 y_{n-1}^{m-1}(f_{n-1}^{m-1} y_{n-1}^{m-1}) \to_t^0 y_{n-1}^{m-1}(g_n^{m+2} z_n^m)$. Thus, $y_{n-1}^{m-1}(g_n^m z_n^m) =_t^0 y_{n-1}^{m-1}(g_n^{m+2} z_n^m)$. Therefore, for any z_n^m, $g_n^m z_n^m =_{n-2}^{m+2} g_n^{m+2} z_n^m$. □

Lemma 5.2. g_n^m *is a one-to-one function from* D_n^m *into* D_{n-2}^m.

Proof. We prove by *reductio* that g_n^m is one-to-one. Suppose g_n^m is many-to-one. Then for some x_n^m and y_n^m, $x_n^m \neq_n^m y_n^m$ and $g_n^m x_n^m =_{n-2}^m g_n^m y_n^m$. But, then, for some z_{n-1}^{m-1}, $x_n^m z_{n-1}^{m-1} \neq_t^0 y_n^m z_{n-1}^{m-1}$. Then, by [3], $z_{n-1}^{m-1}(f_{n-1}^{m-1} z_{n-1}^{m-1} \to_{n-2}^m g_n^m x_n^m) \neq_t^0 z_{n-1}^{m-1}(f_{n-1}^{m-1} z_{n-1}^{m-1} \to_{n-2}^m g_n^m y_n^m)$. So $(f_{n-1}^{m-1} z_{n-1}^{m-1} \to_{n-2}^m g_n^m x_n^m) \neq_{n-2}^m (f_{n-1}^{m-1} z_{n-1}^{m-1} \to_{n-2}^m g_n^m y_n^m)$. Thus, $g_n^m x_n^m \neq_{n-2}^m g_n^m y_n^m$, contradicting the assumption. □

Lemma 5.3. g_n^m *is a function from* D_n^m *onto* D_{n-2}^m.

Proof. We prove by *reductio* that g_n^m is an onto function. Suppose that g_n^m is not onto, and that a_{n-2}^m is not a value of g_n^m. We may assume that g_n^{m-2} is total (that is, we assume that m is the smallest number for which g_n^m is not total). By Theorem 5.1,

$g_n^{m-2} \subseteq g_n^m$. So $a_{n-2}^m \notin D_{n-2}^{m-2}$. Thus, a_{n-2}^m must be of the form $(\sim_{n-2}^m b_{n-2}^m)$, $(b_{n-2}^m \wedge_{n-2}^m c_{n-2}^m)$, or $(b_{n-2}^m \vee_{n-2}^m c_{n-2}^m)$. Suppose a_{n-2}^m is of the form $(\sim_{n-2}^m b_{n-2}^m)$; suppose, in particular, $a_{n-2}^m = (\sim_{n-2}^m b_{n-2}^{m-2})$. By hypothesis, b_{n-2}^{m-2} is a value of g_n^{m-2}. But then a_{n-2}^m must be a value of g_n^m, contradicting the hypothesis of the *reductio*. The same consideration applies to the other cases. Therefore, a_{n-2}^m cannot be any of the forms $(\sim_{n-2}^m b_{n-2}^m)$, $(b_{n-2}^m \wedge_{n-2}^m c_{n-2}^m)$, or $(b_{n-2}^m \vee_{n-2}^m c_{n-2}^m)$) – a contradiction. □

Theorem 5.2. *g_n^m is a one-to-one correspondence from D_n^m into D_{n-2}^m.*

Proof. Immediate from Lemmas 5.2 and 5.3.

Theorem 5.3 (Type Reduction Theorem). *There is a one-to-one correspondence from D_n^m into D_{n+2k}^m.*

Proof. Straightforward from Theorem 5.2.

See again Fig. 5.1. What the Type Reduction Theorem says is that the types $(\cdots((\underbrace{\sigma \to t)\to t)\to\cdots)\to t)}_{2n \text{ times}}\cdots)$ and the types $(\cdots((\underbrace{\sigma\to t)\to t)\to\cdots)\to t)}_{2n+1 \text{ times}}\cdots)$ are reducible to the types σ and $\sigma \to t$, respectively, and thus, that they are not really necessary. By introducing ranks to the shadows of types σ and $\sigma \to t$, we can dispense with the infinite hierarchy of types; it can be 'flattened' into the two types at the bottom. In particular, if $\sigma = e$, for any n, the monadic $(2n)$th-order objects are reducible to the zeroth-order objects (individuals) of type e, and the monadic $(2n+1)$th-order objects are reducible to the monadic first-order objects (sets of individuals) of type $e \to t$.

This does not mean there are only finitely many types in the theory because, for instance, such types as $e \to (e \to t)$, $e \to (e \to (e \to t))$, etc., are not reducible to simpler types. Note that t is here considered an atomic type. Thus, the original domain $\{0, 1\}$ of type t itself is D_0^0, and must be expanded into D_0^∞. In the simplest theory, t is the only atomic type; then there will be only two types in the theory, t and $t \to t$. Simple as it may be, this is a very interesting theory; but we cannot get into details in this paper, except touching upon it a little in the final section.

Another issue we cannot explore in this paper is the relation between shadow theory and classical type theory, set forth by Parigot [8] and others. In the standard, non-classical simple type theory, there are only two type inference rules – (λ-) abstraction and application:

$$\frac{\Gamma, X : \sigma \Rightarrow B : \tau}{\Gamma \Rightarrow (\lambda X.B) : (\sigma \to \tau)} \ (\lambda\text{-})\text{abstraction}$$

$$\frac{\Gamma \Rightarrow P : (\sigma \to \tau) \qquad \Gamma \Rightarrow A : \sigma}{\Gamma \Rightarrow (PA) : \tau} \ \text{application}$$

(λ-)abstraction asserts that if you obtain an expression (or a 'term') B of type τ from the assumption, along with a set Γ of other assumptions, that the variable X it contains is to be filled in with an expression of type σ, then the expression without the filling, designated as $\lambda X.B$, is of type $\sigma \to \tau$. Application asserts that if you obtain the expressions P of type $\sigma \to \tau$ and A of type σ from a set Γ of assumptions, then based on the same assumptions you can apply P to A (or fill in P with A) and obtain an expression of type τ, designated here as PA. These two type inference rules, i.e., if $\sigma \Rightarrow \tau$ then $\Rightarrow \sigma \to \tau$ and $\sigma \to \tau, \sigma \Rightarrow \tau$, constitute the $\to$-fragment of minimal logic. However, the type reduction rule suggested above, $(\sigma \to t) \to t \Rightarrow \sigma$, if we identify t with absurdity $\perp$, is nothing but the classical absurdity rule, $(\sigma \to \perp) \to \perp \Rightarrow \sigma$, or the double-negation elimination rule, $\sim\sim\sigma \Rightarrow \sigma$, if we abbreviate $\sigma \to \perp$ as $\sim\sigma$. The addition of this rule to the two standard type inference rules would make the logic of types classical. (The general approach that treats types just as propositions and type inference rules just as logical inferences is often associated with the terms *the Curry-Howard correspondence* and *propositions-as-types*; see, e.g., Sørensen and Urzyczyn [12].) Parigot and others proposed rules similar to the above to construct a classical type theory, and there seems to be an apparent connection between their theory and shadow theory. But discussion of this topic is beyond the scope of this paper.

What we shall do in the rest of the paper, however, is to substitute $\perp$ for t as the name of the type of truth values and abbreviate $\sigma \to \perp$ as $\sim\sigma$. This will result in much simplification in the following presentation.

5.4 Deductive System *ST*

In this section, we shall discuss a formal deductive system, *ST*, which describes the bottom types σ and $\sim\sigma (= \sigma \to \perp)$ of shadows spelled out in the last section. *ST* is a sequent calculus. The basic understanding of sequent calculi is presupposed.

Language of *ST*. The types in *ST* are defined in abstract grammar as follows:

$$\tau ::= e \mid \perp \mid (\tau \to \tau).$$

For any type τ, $\sim\tau =_{df} (\tau \to \perp)$, $(\sigma \wedge \tau) =_{df} \sim(\sigma \to \sim\tau)$, and $(\sigma \vee \tau) =_{df} (\sim\sigma \to \tau)$.

We divide all types into type 0 and type 1. A type is type 0 if it is prefixed with an even number (including zero) of $\sim$'s, i.e., it is of the form $\underbrace{\sim \ldots \sim}_{2n} t$, whereas a type is a type 1 if it is prefixed with an odd number of $\sim$'s, i.e., it is of the form $\underbrace{\sim \ldots \sim}_{2n+1} t$.

The terms in *ST* are defined as follows, where σ is a type and m and n are non-negative integers:

- For any upper-case alphabet X, X^{2n}_{σ} is a term if σ is of type 0; X^{2n+1}_{σ} is a term if σ is of type 1.
- If A^m_{σ} is a term, then $(\sim A^m_{\sigma})^{m+2k}_{\sigma}$ is a term, where $0 \le k$ if $0 < m$; $0 < k$ if $m = 0$.

- If A^m_σ and B^n_σ are terms, then $(A^m_\sigma \wedge B^n_\sigma)^{max(m,n)+2k}_\sigma$, $(A^m_\sigma \vee B^n_\sigma)^{max(m,n)+2k}_\sigma$, and $(A^m_\sigma \rightarrow B^n_\sigma)^{max(m,n)+2k}_\sigma$ are terms, where $0 \leq k$ if $0 < max(m,n)$; $0 < k$ if $max(m,n) = 0$.
- If $A^m_{\sim\sigma}$ is a term, then $(\forall A^m_{\sim\sigma})^{m-1+2k}_\sigma$ and $(\exists A^m_{\sim\sigma})^{m-1+2k}_\sigma$ are terms, where $0 \leq k$. These read, informally, 'every A' and 'some A', respectively.
- If $A^m_{\sim\sigma}$ and B^n_σ are terms, then $(A^m_{\sim\sigma} B^n_\sigma)^0_\perp$ and $(B^n_\sigma A^m_{\sim\sigma})^0_\perp$ are terms.
- Nothing else is a term.

Deduction rules of *ST*. The deduction rules of *ST* are as follows, where Γ and Δ are finite sets of terms (not sequences; so the usual structural rules can be dispensed with). σ is any type. $0 \leq i, j, m, n$. $0 \leq k$ in *Ax* or in the other rules if the sum involving k is larger than 1; otherwise, $0 < k$.

$$\frac{}{\Gamma, A^0_\perp \Rightarrow A^0_\perp, \Delta}\ Ax$$

$$\frac{\Gamma, (A^m_{\sim\sigma} B^n_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma, (B^n_\sigma A^m_{\sim\sigma})^0_\perp \Rightarrow \Delta}\ SwL$$

$$\frac{\Gamma \Rightarrow (A^m_{\sim\sigma} B^n_\sigma)^0_\perp, \Delta}{\Gamma \Rightarrow (B^n_\sigma A^m_{\sim\sigma})^0_\perp, \Delta}\ SwR$$

$$\frac{\Gamma \Rightarrow (Q^j_{\sim\sigma} A^m_\sigma)^0_\perp, \Delta}{\Gamma (Q^j_{\sim\sigma} (\sim A^m_\sigma)^{m+2k}_\sigma)^0_\perp \Rightarrow \Delta}\ {\sim}L$$

$$\frac{\Gamma, (Q^j_{\sim\sigma} A^m_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma \Rightarrow (Q^j_{\sim\sigma} (\sim A^m_\sigma)^{m+2k}_\sigma)^0_\perp, \Delta}\ {\sim}R$$

In the last two rules, $j < m + 2k$.

$$\frac{\Gamma, (Q^j_{\sim\sigma} A^m_\sigma)^0_\perp, (Q^j_{\sim\sigma} B^n_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma, (Q^j_{\sim\sigma} (A^m_\sigma \wedge B^n_\sigma)^{max(m,n)+2k}_\sigma)^0_\perp \Rightarrow \Delta}\ \wedge L$$

$$\frac{\Gamma \Rightarrow (Q^j_{\sim\sigma} A^m_\sigma)^0_\perp, \Delta \qquad \Gamma \Rightarrow (Q^j_{\sim\sigma} B^n_\sigma)^0_\perp, \Delta}{\Gamma \Rightarrow (Q^j_{\sim\sigma} (A^m_\sigma \wedge B^n_\sigma)^{max(m,n)+2k}_\sigma)^0_\perp, \Delta}\ \wedge R$$

$$\frac{\Gamma, (Q^j_{\sim\sigma} A^m_\sigma)^0_\perp \Rightarrow \Delta \qquad \Gamma, (Q^j_{\sim\sigma} B^n_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma, (Q^j_{\sim\sigma} (A^m_\sigma \vee B^n_\sigma)^{max(m,n)+2k}_\sigma)^0_\perp \Rightarrow \Delta}\ \vee L$$

$$\frac{\Gamma \Rightarrow (Q^j_{\sim\sigma} A^m_\sigma)^0_\perp, (Q^j_{\sim\sigma} B^n_\sigma)^0_\perp, \Delta}{\Gamma \Rightarrow (Q^j_{\sim\sigma} (A^m_\sigma \vee B^n_\sigma)^{max(m,n)+2k}_\sigma)^0_\perp, \Delta}\ \vee R$$

$$\frac{\Gamma \Rightarrow (Q^j_{\sim\sigma}A^m_\sigma)^0_\perp, \Delta \qquad \Gamma, (Q^j_{\sim\sigma}B^n_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma, (Q^j_{\sim\sigma}(A^m_\sigma \to B^n_\sigma)^{max(m,n)+2k}_\sigma)^0_\perp \Rightarrow \Delta} \to L$$

$$\frac{\Gamma, (Q^j_{\sim\sigma}A^m_\sigma)^0_\perp \Rightarrow (Q^j_{\sim\sigma}B^n_\sigma)^0_\perp, \Delta}{\Gamma \Rightarrow (Q^j_{\sim\sigma}(A^m_\sigma \to B^n_\sigma)^{max(m,n)+2k}_\sigma)^0_\perp, \Delta} \to R$$

In the last six rules, $j < max(m, n) + 2k$.

$$\frac{\Gamma \Rightarrow (P^i_{\sim\sigma}A^m_\sigma)^0_\perp, \Delta \qquad \Gamma, (Q^j_{\sim\sigma}A^m_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma, (Q^j_{\sim\sigma}(\forall P^i_{\sim\sigma})^{m+2k}_\sigma)^0_\perp \Rightarrow \Delta} \forall L$$

$$\frac{\Gamma, (P^i_{\sim\sigma}A^m_\sigma)^0_\perp \Rightarrow (Q^j_{\sim\sigma}A^m_\sigma)^0_\perp, \Delta}{\Gamma \Rightarrow (Q^j_{\sim\sigma}(\forall P^i_{\sim\sigma})^{m+2k}_\sigma)^0_\perp, \Delta} \forall R$$

A^m_σ is not in the lower sequent.

$$\frac{\Gamma, (P^i_{\sim\sigma}A^m_\sigma)^0_\perp, (Q^j_{\sim\sigma}A^m_\sigma)^0_\perp \Rightarrow \Delta}{\Gamma, (Q^j_{\sim\sigma}(\exists P^i_{\sim\sigma})^{m+2k}_\sigma)^0_\perp \Rightarrow \Delta} \exists L$$

A^m_σ is not in the lower sequent.

$$\frac{\Gamma \Rightarrow (P^i_{\sim\sigma}A^m_\sigma)^0_\perp, \Delta \qquad \Gamma \Rightarrow (Q^j_{\sim\sigma}A^m_\sigma)^0_\perp, \Delta}{\Gamma \Rightarrow (Q^j_{\sim\sigma}(\exists P^i_{\sim\sigma})^{m+2k}_\sigma)^0_\perp, \Delta} \exists R$$

In the last four rules, $max(i, j) - 2 < m + 2k$.

Each line of the form $\Gamma \Rightarrow \Delta$ is called a *sequent*. $[\![\Gamma \Rightarrow \Delta]\!]$ = T if and only if the conjunction of the members of Γ implies the disjunction of the members of Δ.

The deductive rules so far given do not include sentential connectives in the ordinary sense of the term. (They include shadows of sentences, but they distribute only over the objects of type $\sim\perp(= \perp \to \perp)$.). For sentential connectives, the following rules may be added, although we will not use them in the examples to follow.

$$\frac{}{\Gamma, A^m_\perp \Rightarrow A^{m+2k}_\perp, \Delta} Ax+$$

$$\frac{\Gamma \Rightarrow A^m_\perp, \Delta}{\Gamma, (\sim A^m_\perp)^{m+2k}_\perp \Rightarrow \Delta} P{\sim}L$$

$$\frac{\Gamma, A^m_\perp \Rightarrow \Delta}{\Gamma \Rightarrow (\sim A^m_\perp)^{m+2k}_\perp, \Delta} P{\sim}R$$

$$\frac{\Gamma, A^m_\perp, B^n_\perp \Rightarrow \Delta}{\Gamma, (A^m_\perp \wedge B^n_\perp)^{max(m,n)+2k}_\perp \Rightarrow \Delta}\ P{\wedge}L$$

$$\frac{\Gamma \Rightarrow A^m_\perp, \Delta \qquad \Gamma \Rightarrow B^n_\perp, \Delta}{\Gamma \Rightarrow (A^m_\perp \wedge B^n_\perp)^{max(m,n)+2k}, \Delta}\ P{\wedge}R$$

$$\frac{\Gamma, A^m_\perp \Rightarrow \Delta \qquad \Gamma, B^n_\perp \Rightarrow \Delta}{\Gamma, (A^m_\perp \vee B^n_\perp)^{max(m,n)+2k}_\perp \Rightarrow \Delta}\ P{\vee}L$$

$$\frac{\Gamma \Rightarrow A^m_\perp, B^n_\perp, \Delta}{\Gamma \Rightarrow (A^m_\perp \vee B^n_\perp)^{max(m,n)+2k}_\perp, \Delta}\ P{\vee}R$$

$$\frac{\Gamma \Rightarrow A^m_\perp, \Delta \qquad \Gamma, B^n_\perp \Rightarrow \Delta}{\Gamma, (A^m_\perp \to B^n_\perp)^{max(m,n)+2k}_\perp \Rightarrow \Delta}\ P{\to}L$$

$$\frac{\Gamma, A^m_\perp \Rightarrow B^n_\perp, \Delta}{\Gamma \Rightarrow (A^m_\perp \to B^n_\perp)^{max(m,n)+2k}_\perp, \Delta}\ P{\to}R$$

In most cases the rank of shadows involved above should be 0. If it is, the above rules may be considered special instances of the rules for the term connectives spelled out before, where the 'truth predicate' ('…is true') $Q^1_{\sim\perp}$, denoting the identity function ($0 \mapsto 0$; $1 \mapsto 1$), may be taken to be implicitly appended in front of each sentence.

There are other rules that can be included; obvious instances are the application and (λ-)abstraction rules. However, for the formalization of most natural language inferences, the above rules seem sufficient.

At this point, it would be customary to give the intended semantics to the deduction system just presented; however, in this case the language and the rules of *ST* so closely match shadow theory given in the last section that its semantics is self-evident. Note that $\wedge$, $\vee$, and $\sim$ are used in the last section as glb, lub, and complement and in this section as conjunction, disjunction, and negation, but that they correspond to one another. *SwL* and *SwR* are necessary because we stipulate that only the second term can distribute over the first term in a combination of terms, and not the other way around.

Examples. The content of *ST* is probably best understood by its applications to inferences in natural language. Thus, we shall consider a few such applications. Let us first compare the following two inferences, which involve compound NPs:

41. Adam is male. Betty is female. Therefore, Adam and Betty are male or female.
42. Adam is male. So is Bob. Therefore, Adam and Bob are both male or both female.

In (41), *and* has a wider scope than *or*, whereas in (42), *or* has a wider scope. The final conclusion of (41) follows from the intermediate conclusions, *Adam is male or* *female* and *Betty is male or female*; the final conclusion of (42) follows from *Adam and Bob are both male*. The following are the derivations in *ST* of (41) and (42), respectively:

$$\dfrac{\dfrac{\dfrac{\dfrac{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow (M^1_{\sim e}A^0_e)^0_\bot, (A^0_e F^1_{\sim e})^0_\bot}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow (A^0_e M^1_{\sim e})^0_\bot, (A^0_e F^1_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow (A^0_e (M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow ((M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e} A^0_e)^0_\bot} \quad \dfrac{\dfrac{\dfrac{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow (B^0_e M^1_{\sim e})^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow (B^0_e M^1_{\sim e})^0_\bot, (B^0_e F^1_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow (B^0_e (M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow ((M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e} B^0_e)^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (F^1_{\sim e}B^0_e)^0_\bot \Rightarrow ((M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e} (A^0_e \wedge B^0_e)^4_e)^0_\bot}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{(M^1_{\sim e}A^0_e)^0_\bot, (M^1_{\sim e}B^0_e)^0_\bot \Rightarrow (M^1_{\sim e}A^0_e)^0_\bot, ((A^0_e \wedge B^0_e)^2_e F^1_{\sim e})^0_\bot \quad (M^1_{\sim e}A^0_e)^0_\bot, (M^1_{\sim e}B^0_e)^0_\bot \Rightarrow (M^1_{\sim e}B^0_e)^0_\bot, ((A^0_e \wedge B^0_e)^2_e F^1_{\sim e})^0_\bot}{(M^1_{\sim e}A^0_e)^0_\bot, (M^1_{\sim e}B^0_e)^0_\bot \Rightarrow (M^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\bot, ((A^0_e \wedge B^0_e)^2_e F^1_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (M^1_{\sim e}B^0_e)^0_\bot \Rightarrow ((A^0_e \wedge B^0_e)^2_e M^1_{\sim e})^0_\bot, ((A^0_e \wedge B^0_e)^2_e F^1_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (M^1_{\sim e}B^0_e)^0_\bot \Rightarrow ((A^0_e \wedge B^0_e)^2_e (M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e})^0_\bot}}{(M^1_{\sim e}A^0_e)^0_\bot, (M^1_{\sim e}B^0_e)^0_\bot \Rightarrow ((M^1_{\sim e} \vee F^1_{\sim e})^3_{\sim e} (A^0_e \wedge B^0_e)^2_e)^0_\bot}$$

As you can see in the final lines, $(A \wedge B)$ is at rank 2 in the second derivation, but it is at rank 4 in the first derivation; $(M \vee F)$ is at rank 3 in both derivations, so it is distributive (or has a scope) over $(A \wedge B)$ in the second derivation, but $(A \wedge B)$ is distributive over it in the first derivation. The rules given above allow us to assign different ranks, but the rank of $(A \wedge B)$ in the second derivation must always be lower, and that in the first derivation higher, than the rank of $(M \vee F)$. $(A \wedge B)$ is of type e, and $(M \vee F)$ of type $\sim e$, throughout both derivations.

Compare these with the following derivations of (32) and (33), which employ the standard, non-classical type theory:

$$\dfrac{\dfrac{\dfrac{(M_{\sim e}A_e)_\bot, (F_{\sim e}B_e)_\bot \Rightarrow (M_{\sim e}A_e)_\bot, (F_{\sim e}A_e)_\bot}{(M_{\sim e}A_e)_\bot, (F_{\sim e}B_e)_\bot \Rightarrow ((M_{\sim e} \vee F_{\sim e})_{\sim e}A_e)_\bot}}{(M_{\sim e}A_e)_\bot, (F_{\sim e}B_e)_\bot \Rightarrow (A \uparrow_{\sim\sim e} (M_{\sim e} \vee F_{\sim e})_{\sim e})_\bot} \quad \dfrac{\dfrac{(M_{\sim e}A_e)_\bot, (F_{\sim e}B_e)_\bot \Rightarrow (M_{\sim e}B_e)_\bot, (F_{\sim e}B_e)_\bot}{(M_{\sim e}A_e)_\bot, (F_{\sim e}B_e)_\bot \Rightarrow ((M_{\sim e} \vee F_{\sim e})_{\sim e}B_e)_\bot}}{(M_{\sim e}A_e)_\bot, (F_{\sim e}B_e)_\bot \Rightarrow (B \uparrow_{\sim\sim e} (M_{\sim e} \vee F_{\sim e})_{\sim e})_\bot}}{((A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e}(M_{\sim e} \vee F_{\sim e})_{\sim e})_\bot}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow (M_{\sim e}A_e)_\bot, (F \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot}{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow (A \uparrow_{\sim\sim e} M_{\sim e})_\bot, (F \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot} \quad \dfrac{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow (M_{\sim e}B_e)_\bot, (F \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot}{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow (B \uparrow_{\sim\sim e} M_{\sim e})_\bot, (F \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot}}{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow ((A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e} M_{\sim e})_\bot, (F \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot}}{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow (M \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot, (F \uparrow_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot}}{(M_{\sim e}A_e)_\bot, (M_{\sim e}B_e)_\bot \Rightarrow ((M \uparrow_{\sim\sim\sim e} \vee F \uparrow_{\sim\sim\sim e})_{\sim\sim\sim e} (A \uparrow_{\sim\sim e} \wedge B \uparrow_{\sim\sim e})_{\sim\sim e})_\bot}$$

Here $X \uparrow$ is the *type-lifting* of X. Even though these derivations may look simpler than the previous derivations (especially because they do not have superscripts for ranks), they involve type-lifting, which is a conceptually complicated process. Basically, if X of type t denotes a member of D^m_σ, $X \uparrow$ will be of type $\sim\sim t$ and denotes the corresponding object in $D^{m-2}_{\sim\sim\sigma}$. This mechanism allows the relevant compound expression involving $X \uparrow$ to have a wider scope. In contrast, in *ST* no expressions, denotations, or types need to be changed.

One might say that in *ST* the ranks simply play the roles of higher types, so nothing is gained. While it is true that we can eliminate the relevant higher types thanks to the roles the ranks play, it is not true that nothing is gained as a result.

Objects at lower ranks are also at higher ranks (i.e., $D^m_\sigma \subseteq D^{m+2}_\sigma$), so we do not need to switch denotations at lower ranks when we move up to higher ranks. In contrast, in the semantics based on the standard, non-classical type theory, we must switch denotations when we move up.

ST can also deal with *non-distributive conjunctions*. Consider the following example:

43. Adam and Betty together carried a piano upstairs. Adam and Betty were each paid \$50. Therefore, Adam and Betty together carried a piano upstairs and were each paid \$50.

Adam and Betty in the first premise is a non-distributive conjunction: we cannot derive from the premise *Adam carried a piano upstairs and Betty carried a piano upstairs*. *Adam and Betty* in the second premise, on the other hand, is distributive: the premise is equivalent to *Adam was paid \$50 and Betty was paid \$50*. The following is the derivation of (43) in *ST*:

$$\dfrac{\dfrac{\dfrac{(C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp, (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp \Rightarrow (C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp}{(C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp, (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp \Rightarrow ((A^0_e \wedge B^0_e)^2_e C^3_{\sim e})^0_\perp} \qquad \dfrac{(C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp, (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp \Rightarrow (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp}{(C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp, (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp \Rightarrow ((A^0_e \wedge B^0_e)^2_e P^1_{\sim e})^0_\perp}}{(C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp, (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp \Rightarrow ((A^0_e \wedge B^0_e)^2_e (C^3_{\sim e} \wedge P^1_{\sim e})^3_{\sim e})^0_\perp}}{(C^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp, (P^1_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp \Rightarrow ((C^3_{\sim e} \wedge P^1_{\sim e})^3_{\sim e}(A^0_e \wedge B^0_e)^2_e)^0_\perp}$$

The key here is the distribution (or scope) relations among $P^1_{\sim e}$, $C^3_{\sim e}$, $(C \wedge P)^3_{\sim e}$ and $(A \wedge B)^2_e : P^1_{\sim e} < (A \wedge B)^2_e < C^3_{\sim e}, (C \wedge P)^3_{\sim e}$. Here $C^3_{\sim e}$ must denote an object properly in $D^3_{\sim e}$, i.e., in $D^3_{\sim e}$ but not in $D^1_{\sim e}$. Compare this example with (44) below:

44. Adam carried a piano upstairs. Betty also carried a piano (possibly a different one) upstairs. Therefore, Adam and Betty each carried one piano upstairs.

The derivation of this inference is as follows:

$$\dfrac{(C^3_{\sim e}A^0_e)^0_\perp, (C^3_{\sim e}B^0_e)^0_\perp \Rightarrow (C^3_{\sim e}A^0_e)^0_\perp \qquad (C^3_{\sim e}A^0_e)^0_\perp, (C^3_{\sim e}B^0_e)^0_\perp \Rightarrow (C^3_{\sim e}B^0_e)^0_\perp}{(C^3_{\sim e}A^0_e)^0_\perp, (C^3_{\sim e}B^0_e)^0_\perp \Rightarrow (C^3_{\sim e}(A^0_e \wedge B^0_e)^4_e)^0_\perp}$$

Assuming that C (*carry a piano upstairs*) is the same as before (i.e., $C^3_{\sim e}$), $(A \wedge B)$ in the conclusion must be at rank 4 or higher in order to distribute over C.

We conclude this section by considering some examples that involve quantifiers. Put simply, a universal quantifier is to be understood as denoting a (possibly infinite) conjunctive shadow, and an existential quantifier is to be understood as denoting a (possibly infinite) disjunctive shadow. Thus understood, quantifiers are conceptually nothing but a simple extension of conjunctive and disjunctive NPs. Consider

45. Every male person is male.
46. Adam is a person. Adam is male. Therefore, someone is male.
47. Pegasus is a winged horse. Pegasus exists. Therefore, a winged horse exists.

These inferences are formalized respectively as follows:

$$\begin{array}{c}
(M^1_{\sim e}A^0_e)^0_\perp, (P^1_{\sim e}A^0_e)^0_\perp \Rightarrow (M^1_{\sim e}A^0_e)^0_\perp \\
\hline
(A^0_e M^1_{\sim e})^0_\perp, (A^0_e P^1_{\sim e})^0_\perp \Rightarrow (M^1_{\sim e}A^0_e)^0_\perp \\
\hline
(A^0_e (M^1_{\sim e} \wedge P^1_{\sim e})^3_{\sim e})^0_\perp \Rightarrow (M^1_{\sim e}A^0_e)^0_\perp \\
\hline
((M^1_{\sim e} \wedge P^1_{\sim e})^3_{\sim e} A^0_e)^0_\perp \Rightarrow (M^1_{\sim e}A^0_e)^0_\perp \\
\hline
\Rightarrow (M^1_{\sim e}(\forall (M^1_{\sim e} \wedge P^1_{\sim e})^3_{\sim e})^2_e)^0_\perp
\end{array}$$

$$\frac{(P^1_{\sim e}A^0_e)^0_\perp, (M^1_{\sim e}A^0_e)^0_\perp \Rightarrow (P^1_{\sim e}A^0_e)^0_\perp \qquad (P^1_{\sim e}A^0_e)^0_\perp, (M^1_{\sim e}A^0_e)^0_\perp \Rightarrow (M^1_{\sim e}A^0_e)^0_\perp}{(P^1_{\sim e}A^0_e)^0_\perp, (M^1_{\sim e}A^0_e)^0_\perp \Rightarrow (M^1_{\sim e}(\exists P^1_{\sim e})^2_e)^0_\perp}$$

$$\frac{((W^1_{\sim e} \wedge H^1_{\sim e})^3_{\sim e}P^0_e)^0_\perp, (E^1_{\sim e}P^0_e)^0_\perp \Rightarrow ((W^1_{\sim e} \wedge H^1_{\sim e})^3_{\sim e}P^0_e)^0_\perp \qquad ((W^1_{\sim e} \wedge H^1_{\sim e})^3_{\sim e}P^0_e)^0_\perp, (E^1_{\sim e}P^0_e)^0_\perp \Rightarrow (E^1_{\sim e}P^0_e)^0_\perp}{((W^1_{\sim e} \wedge H^1_{\sim e})^3_{\sim e}P^0_e)^0_\perp, (E^1_{\sim e}P^0_e)^0_\perp \Rightarrow (E^1_{\sim e}(\exists (W^1_{\sim e} \wedge H^1_{\sim e})^3_{\sim e})^2_e)^0_\perp}$$

Note that in the last derivation, *exists* is not a quantifier but a predicate that denotes a first-order property just like *is winged* and *is a horse* do. This is in harmony with the fact that, grammatically, they are of the same kind.

As you can see in the above examples, *ST*, with the help of its rank differences, has a mechanism to deal with various scope relations without resorting to type-lifting. In this section we have fastidiously indexed all the formulas in the derivations, and that might give you the impression that *ST* is unnecessarily complex, but clearly many indices may be omitted in practice, and when they are, *ST* is no more complex than the non-classically typed systems.

5.5 Conclusion

This paper set forth shadow theory. According to shadow theory, every object of some type at some rank, say $n(\neq 0)$, has its negative shadows of the same type at the ranks n or higher, and every set of objects of the same type at the ranks no higher than n has its conjunctive and disjunctive shadows of the same type at the ranks n or higher. There are infinitely many ranks, and the objects at each rank constitute a Boolean algebra. This may initially seem to require an intolerably large ontology, but, in fact, as a result of the rank distinction, all higher types in the infinite hierarchy are reducible to the bottom two types, resulting in a simpler ontology. Many NPs in natural languages like English, including many quantifiers, can be considered to denote shadows, and various scope differences can be appropriately treated with the rank distinction. In sum, shadow theory is a very attractive, promising ontological theory that serves well for the semantics of natural language.

Finally, the types of truth values $\perp$ and $\sim\perp$ may deserve special attention. More precisely, the former is the type of truth values and the latter is the type of functions from the truth values to the truth values, or the type of truth functions. $D^0_\perp = D^0_0 = \{0, 1\}$, and $D^1_{\sim\perp} = D^1_1$ consists of four functions – $0 \mapsto 0; 1 \mapsto 1$ (identity function),

$0 \mapsto 1; 1 \mapsto 0$ (negation function), and two constant functions, $0, 1 \mapsto 0$ and $0, 1 \mapsto 1$. The first function may also be considered the truth predicate for D_0^0. $D_\perp^2 = D_0^2$ consists of $2^4 = 16$ functions from those four functions to $D_\perp^0 = \{0, 1\}$, and $D_{\sim\perp}^3 = D_1^3$ consists of 2^{16} functions, and so on. Each D_1^{2n+1} contains the truth predicate for D_0^{2n}. These truth predicates constitute an infinite hierarchy of truth predicates for the ever expanding domain for the type $\perp$, in a way similar to Tarski's [13] well-known hierarchy. Also, if we identify 0 with $\emptyset$ and 1 with $\{\emptyset\}$, then D_1^1 may be considered the set of (the characteristic functions of) the sets consisting of $\emptyset$ and $\{\emptyset\}$, and D_0^2 as the set of sets of sets, and so on. Thus, the ranks in shadow theory correspond to the ranks in set theory, except that the set-theoretic rank = the shadow-theoretic rank +1.

In this paper, we have generally taken the members of D_σ^m to be the functions from $D_{\sim\sigma}^{m-1}$ to $D_\perp^0$, but we can construct theories that replace $D_\perp^0$ with $D_\perp^{2m}$ or $D_{\sim\perp}^{2m-1}$ for various m's (> 0) – that is, theories with more than two truth values. As is explained in Akiba [2], such theories introduce intensionality without employing possible worlds, and can also be used to deal with vagueness in language. Further research is called for also in this direction.

References

1. Akiba, K. 2009. A new theory of quantifiers and term connectives. *Journal of Logic, Language and Information* 18: 403–431.
2. Akiba, K. 2014. Boolean-valued sets as vague sets. In *Vague objects and vague identity*, ed. K. Akiba and A. Abasnezhad, 175–195. Dordrecht: Springer.
3. Barwise, J., and R. Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4: 159–219.
4. Church, A. 1940. A formulation of the simple theory of types. *Journal of Symbolic Logic* 5: 56–68.
5. Kripke, S. 1980. *Naming and necessity*. Cambridge: Harvard University Press.
6. Lewis, D. 1970. General semantics. *Synthese* 22: 18–67.
7. Montague, R. 1973. The proper treatment of quantification in ordinary English. In *Approaches to natural language*, ed. J. Hintikka, J. Moravcsik, and P. Suppes, 221–242. Dordrecht: Reidel.
8. Parigot, M. 1992. $\lambda\mu$-calculus: an algorithmic interpretation of classical natural deduction. In *Logic programming and automated reasoning: international conference, LPAR '92*. Lecture notes in artificial intelligence, vol. 624, ed. A. Voronkov, 190–201. Dordrecht: Springer.
9. Partee, B. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. J. Groenendijik, D. de Jongh, and M. Stokhof, 115–143. Dordrecht: Foris.
10. Partee, B., and M. Rooth. 1983. Generalized conjunction and type ambiguity. In *Meaning, use and interpretation of language*, ed. R. Bäuerle, C. Schwarze, and A. von Stechow, 361–383. Berlin: Walter de Gruyter.
11. Russell, B. 1905. On denoting. *Mind* 14: 479–493.
12. Sørensen, M.H., and P. Urzyczyn. 2006. *Lectures on the Curry-Howard isomorphism*. Amsterdam: Elsevier.
13. Tarski, A. 1956. The concept of truth in formalized languages. In *Logic, semantics, metamathematics*, ed. A. Tarski, 152–278. Oxford: Oxford University Press.

Chapter 6
Quantifiers and Referential Use

Mario Gómez-Torrente

Abstract Referential uses of quantified determiner phrases other than descriptions have not been extensively considered. In this paper they are considered in some detail, and related to referential uses of descriptions. The first aim is to develop the observation that, contrary to the currently received view that it is only for descriptions that referential uses are frequent and standard, arising in run-of-the-mill contextual scenarios, this is in fact the case for all usual kinds of quantifier phrases. A second aim is to offer a preliminary discussion of how these data about quantifier phrases other than descriptions constrain the feasible extensions of theories of descriptions to cover the referential uses of quantifier phrases in general. I argue that the data don't support a semantic explanation of referential uses of descriptions, and in fact suggest problems for several semantic theories of referential uses of quantifier phrases in general. I also argue that pragmatic theories of referential uses of quantifier phrases in general might plausibly explain standard referential uses as involving a genus of particularized conversational implicatures in which no conversational maxims are "flouted" or even violated, rather than generalized implicatures or particularized implicatures of Grice's "exploitative" type. I nevertheless emphasize that I don't take the dispute between semantic and pragmatic theories of referential use to have been satisfactorily resolved.

The referential uses of descriptions (both definite and indefinite) have received a great deal of scholarly attention. As a result, our knowledge of the linguistic data relevant to the study of such uses, as well as our grasp of the range of theories that may account for them, can be said to be fairly extensive. Referential uses of other quantified determiner phrases, by contrast, have certainly not been extensively

M. Gómez-Torrente (✉)
Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, México D.F. 04510, Mexico
e-mail: mariogt@unam.mx

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_6

considered by any means. This paper is a brief essay on referential uses of quantified determiner phrases in general. A number of mistaken beliefs seem to surround these uses in the literature that does consider them, including some misconceptions of the basic properties they exhibit. In particular, the currently received view appears to be that at most only referential uses of definite (and possibly indefinite) descriptions are frequent and standard, not requiring *recherché* contextual scenarios. The paper's first aim, pursued in Sect. 6.1, is to develop with minimal detail the observation that, contrary to the received view, for all typical kinds of quantified determiner phrases (quantifier phrases, for short), referential uses are frequent and can be perfectly standard, arising in run-of-the-mill contextual scenarios.

A second aim of the paper, pursued in Sect. 6.2, is to offer a preliminary, necessarily sketchy discussion of how these data constrain the feasible extensions of theories of descriptions to cover the referential uses of quantifier phrases in general. The received view that referential uses of quantifier phrases other than descriptions are different from referential uses of descriptions in the way just mentioned has lately provided great stimulus to views on which these latter uses, and especially those of definite descriptions, ought to receive a thoroughly semantic explanation. There are two reasons for this. First, if the two kinds of uses are substantively different in the mentioned way, the road is open for the suggestion that, even if the semantics of other quantifiers is strictly non-referential, the frequency and standardness of the referential uses of the definite (and possibly the indefinite) article reflect an appropriate semantic convention. Second, the purported differences suggest that while a standard pragmatic explanation of referential uses of quantifier phrases other than descriptions is required, such an explanation is unavailable in the case of referential uses of descriptions. My own view, motivated to a great extent by the observations of Sect. 6.1, is that the data appear to be more complex than usually thought, and are far from unequivocally supporting a semantic explanation of referential uses of descriptions. Another motivation is my impression that fully pragmatic theories have at their disposal a simple apparatus of rarely heeded concepts and distinctions that can reasonably account for the standard, non-*recherché* character of many referential uses of quantifier phrases in general (including descriptions). In particular, a suggestion that emerges from the discussion in Sect. 6.2 is that pragmatic theories might plausibly explain these uses as involving a genus of particularized conversational implicatures in which no conversational maxims are "flouted" or even violated, rather than generalized implicatures or particularized implicatures of Grice's "exploitative" type.

In the final Sect. 6.3 I offer some brief comments on the question of whether or to what extent the preceding considerations favor any particular kind of theory of the referential uses of quantifier phrases over the others. I emphasize that I don't take the dispute between semantic and pragmatic theories of referential use to have been satisfactorily resolved as a result of the preceding discussion.

6.1 The Referential Uses of Quantifier Phrases

It is widely agreed that definite descriptions have both what Keith Donnellan [7] called "referential" uses and what he called "attributive" uses. There is less than wide agreement as to what would be an exact, counterexample-free, theoretical characterization of the distinction. The rough idea, not intended to be free from counterexamples, is that with a referential use of a definite description *the F* a speaker intends to communicate, and typically successfully manages to communicate, a content that in some distinctive sense *involves* some *particular* object (or objects, in plural descriptions) that the speaker means to be the F; with an attributive use of *the F*, by contrast, a speaker intends to communicate, and typically successfully manages to communicate, a *general* content not involving (in the same distinctive sense) any particular object meant to be the F. Note that on this understanding the content of an utterance containing a description *the F* used attributively may in some cases be a not *purely* general content, i.e. may be a content involving some particular objects or others, and in fact even the object that turns out to be the denotation of *the F*. For example, suppose that, in an utterance of

The man among those over there who will get here first will get 10 dollars,

made by someone while pointing at three men she is looking at in plain view, those three men in the domain of quantification come to be involved in the content expressed, as seems reasonable to expect on standard assumptions about demonstratives. Then that content, though general, will not be a *purely* general content. On a Russellian account, e.g., this will be the not purely general content that (say) among *a*, *b* and *c* there is exactly one man who will get here first and will get 10 dollars. Still, this content will not be the content relevant to a referential use of the description "the man among those over there who will get here first", a content which must involve in the desired way the particular man who will get here first (imagine a use in which the speaker is thinking of one of the men in particular because she for some reason knows that he will get here first) and predicate of *him* that he will get 10 dollars.[1]

Vague and hard to pin down as the idea is, most would agree that they know a referential use when they see it. One of Donnellan's examples illustrates the distinction well. Consider the sentence

[1]When talking of the "attributive" use of descriptions, Donnellan had predominantly in mind cases in which the intuitive domain of quantification is highly unrestricted (such as the attributive use of (1) mentioned below in the main text), not cases where (as in the example just discussed) this domain is highly restricted or even restricted to a few objects in plain view. When I speak of attributive or non-referential uses, by contrast, I mean explicitly not to exclude such cases; to insist, the mark of attributivity of a use of *the F* in the sense relevant to the present discussion is merely that the speaker wishes to express a general content and is not thinking of a particular object as being the F.

1. The murderer of Smith is insane.

First think of (1) as uttered by police detective Jones as she comes upon the horribly mutilated dead body of Smith. Jones has no idea of who has murdered Smith, and as she makes her utterance she is not thinking of any particular person (other than Smith; but Jones is not thinking of Smith as his own murderer); she just wants to let it be known that she thinks that whoever killed Smith must be insane. Now suppose Jones and her colleagues arrest Peters and charge him with Smith's murder. Suppose further that Peters is in fact the murderer.[2] And finally suppose that he is standing trial and he frantically moves in the dock. Jones is in the courtroom and makes a new utterance of (1) while looking at Peters, wishing to let it be known that she thinks that Peters is insane. Jones' utterance of (1) in the murder scene contains an attributive, non-referential use of "the murderer of Smith"; her utterance of (1) in the courtroom contains a referential use of that same description.

Indefinite descriptions also give rise both to referential and attributive uses (on the obvious extended understanding of these expressions). Consider the sentence

2. A man wearing big boots is stealing our lemons.

First think of (2) as uttered by me in the presence of my wife as we see our lemon tree thoroughly plundered and the footprints of big boots on the ground by the tree. I have no idea of who is stealing our lemons, and as I make my utterance I am not thinking of any particular person; I just want to let it be known that I think that someone or other who is wearing big boots is stealing our lemons. Now suppose someone else and her husband are watching as someone wearing big boots makes some suspicious moves around the garden where someone has been stealing their lemons. She then makes an utterance of (2) while looking at the man, wishing to let it be known that she thinks that he is stealing their lemons. My utterance of (2) contains an attributive use of "a man wearing big boots"; the woman's utterance of (2) contains a referential use of that same description.

It is also widely agreed that other quantifier phrases have both referential as well as attributive uses (on an obvious extended understanding of these expressions). It is common to cite here some examples from Stephen Neale [15] (adapted in turn from examples by Mark Sainsbury [21] and Martin Davies [4]). The first example concerns "everyone":

---

[2]Donnellan's original discussion relied heavily on cases of referential use of sentences of the form *The F is G* in which *F* is not (uniquely) satisfied by the object the speaker has in mind—cases of *misdescription*. In such cases, Donnellan appears to think that the relevant utterance of *The F is G* can still semantically express a true content, against a Russellian theory on which it would express a false content (and against a Fregean-Strawsonian theory on which it would fail to express a content with a truth value). In most of the recent literature, however, and in fact even in the literature more sympathetic to Donnellan, there seems to be an agreement that this aspect of Donnellan's view is to be rejected: *The F is G* does not semantically express a true content if *F* is not (uniquely) satisfied (possibly in some appropriately restricted domain), regardless of whether the use of *the F* is referential. In line with this tendency, I will take it for granted that *The F is G* never semantically expresses a true content unless *F* is (uniquely) satisfied, and that cases of misdescription are not relevant to the resolution of the debates I will be considering.

> Suppose it is common knowledge that Smith is the only person taking Jones' seminar. One evening, Jones throws a party and Smith is the only person who turns up. A despondent Jones, when asked the next morning whether his party was well attended, says,
>
> 3. Well, everyone taking my seminar turned up
>
> fully intending to inform me that only Smith attended. (pp. 87–8)

Here Jones' utterance of (3) contains a referential use of "everyone taking my seminar", whereby Jones intends to communicate that *Smith* (and he alone) attended his party. (By contrast, imagine that Jones had uttered (3) in a different situation in which his seminar—a *web* seminar, as it happens—had been attended by many people he doesn't really know, and he had thrown an end-of-term party for them that all of them attended, but that neither Jones nor his audience managed to attend; this utterance of (3) would contain an attributive use of "everyone taking my seminar".) The second example concerns the quantifier "most":

> Suppose that Scott Soames, David Lewis, and I are the only three people in Lewis's office. Soames has never played cricket and knows that I know this. In addition, Soames wants to know whether Lewis and I have ever played cricket, so I say
>
> 4. Most people in this room have played cricket
>
> fully intending to communicate to Soames that Lewis and I have both played cricket. (p. 88)

Here Neale's utterance of (4) contains a referential use of "most people in this room", by means of which he intends to communicate to Soames that *Neale* and *Lewis* have played cricket before. (By contrast, imagine that Neale had uttered (4) in a situation in which he had just been given the results of a poll on cricket-playing habits conducted among the numerous occupants of a large room he hasn't even seen; this utterance of (4) would contain an attributive use of "most people in this room".)

It is a common observation, made e.g. by Marga Reimer ([20], 96, 99 n.24), Kent Bach ([2], 226), Michael Devitt ([5], 283), Neale ([16], 173), and Peter Ludlow and Neale ([14], 304), that these examples of referential uses of quantifier phrases other than descriptions are substantively different from usual examples of referential uses of definite (and for some authors, also indefinite) descriptions. Those examples (unlike, say, the example of referential use of (1) above) seem to have an air of *abnormality* or *nonstandardness*. The point is also put in a perhaps more risky way from an empirical point of view, saying that referential uses of quantifier phrases other than descriptions (which are assumed to be relevantly similar to the examples involving (3) and (4)) would seem to be comparatively *infrequent* in real communication (unlike referential uses of descriptions, which are assumed to be predominantly relevantly similar to the example involving (1)).

These observations seem correct, and there are natural reasons why they should be correct. First, the relevant utterances of (3) and (4), unlike the relevant utterance of (1), are able to communicate the intended contents only in virtue of their being embedded in contextual setups where certain assumptions must be common to speaker and audience which will not be straightforwardly recoverable from the physical setup of the conversation, and which will in fact be rather special. Second, in these examples the speaker makes a deliberately quirky utterance given the direction and purposes of the conversation, and in fact an utterance which requires from the audience the actual performance of a slightly sophisticated inference. Thus, in the case of the utterance of (3), it must be common knowledge that Smith is the only person taking Jones' seminar, but this is certainly a rather special piece of information that Jones could not straightforwardly assume to be possessed by a typical interlocutor who joins the conversational setup; in the case of (4), Neale must know that Soames knows that Neale knows that Soames has never played cricket, which is again hardly a piece of information Neale could straightforwardly assume to be possessed by a typical interlocutor who joins the conversation (perhaps from outside Lewis's office). Further, it would seem that if Jones' audience recognizes Jones' intent, this will be because they as a matter of fact carry out the moderately sophisticated inference that the relevant content involving Smith follows from the content of Jones' initially odd utterance (odd, that is, given that he had been asked how many people attended his party) and from the piece of common knowledge about Smith being the only person taking the seminar; similarly, Soames must actually do a moderate bit of inferring from (the content of) Neale's quirky utterance and from his knowledge of Neale's beliefs about him.[3]

From the mentioned observations concerning the common examples of referential uses of quantifier phrases other than descriptions, the mentioned authors invariably conclude the stronger claim that referential uses of quantifier phrases other than descriptions are not frequent and at any rate are abnormal or nonstandard. But I think this conclusion is wrong.

Let's go back to Smith's murder case, but let's imagine that the police investigation developed somewhat differently. Now we are to imagine that Jones and her colleagues arrested seven people, Adams, Barnes, Crane, Daniels, Evans, Foster and Green, and charged all of them with Smith's murder; according to the police, they all acted together and played comparable roles in the brutal slaying, and we can suppose that the police are right. Imagine further that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are now standing trial in the dock, and that Jones is again present in the courtroom. Consider the following sentences:

5. (a) Every murderer of Smith is insane.
   (b) Every guy in the dock is insane.
6. (a) Most murderers of Smith are insane.
   (b) Most guys in the dock are insane
7. (a) Many murderers of Smith are insane.
   (b) Many guys in the dock are insane.
8. (a) Several murderers of Smith are insane.
   (b) Several guys in the dock are insane.

[3] As Neale puts it, "The natural thing to say is that given his background beliefs and given the quantificational proposition expressed by my utterance in the context in question, Soames was able to infer the truth of a particular object-dependent proposition" ([15], 88).

9. (a) Some murderers of Smith are insane.
   (b) Some guys in the dock are insane.
10. (a) A few murderers of Smith are insane.
   (b) A few guys in the dock are insane.

It is of course easy to imagine utterances of (5)–(10) by which an utterer would not be attempting to communicate contents about any particular persons. But I think it's also easy (and I would say *easier*) to see how, if some of the detainees in the dock behave in suitable ways, Jones can use the quantifier phrases in all of these sentences of the form [Q_x: *x is a murderer of Smith*] *x is insane*[4] intending to communicate, and successfully communicating, a variety of contents involving some particular detainees, meaning in each case that those particular detainees are or provide [Q_x : *x is a murderer of Smith*]; and hence it is mandatory to view the corresponding utterances as containing referential uses of the corresponding quantifier phrases.

Imagine first that all the detainees are moving frantically in the dock. Jones may then make an utterance of either (5a) or (5b) intending to communicate, and successfully managing to communicate to an interlocutor sitting next to her in the courtroom, that *Adams, Barnes, Crane, Daniels, Evans, Foster* and *Green* are insane. Jones' utterance of (5a) thus contains a referential use of "every murderer of Smith" and her utterance of (5b) contains a referential use of "every guy in the dock". Second, imagine that Adams, Barnes, Crane, Daniels and Evans are moving frantically in the dock, while Foster and Green are calmly seated. If Jones then makes an utterance of either (6a), (6b), (7a), (7b), (8a) or (8b) intending to communicate that *Adams, Barnes, Crane, Daniels* and *Evans* are insane, she will successfully manage to communicate precisely that to an interlocutor sitting next to her in the courtroom. Jones' utterances of (6a), (6b), (7a), (7b), (8a) or (8b) contain referential uses of "most murderers of Smith", "most guys in the dock", "many murderers of Smith", "many guys in the dock", "several murderers of Smith" and "several guys in the dock", respectively. Finally, imagine that it is Evans, Foster and Green who move frantically in the dock while Adams, Barnes, Crane and Daniels are calmly seated. Jones may make an utterance of either (9a), (9b), (10a) or (10b) intending to communicate, and successfully managing to communicate to an interlocutor sitting next to her in the courtroom, that *Evans, Foster* and *Green* are

[4]In order to make some general claims about referential uses of quantifier phrases here and in what follows, we can assume that the relevant sentences have this or some similar formal representation at some deep level of linguistic structure. Having introduced this notation will also help when we discuss some views of referential uses in Sect. 6.2. As usual, the notation [$Q_x : F$] symbolizes Q as a restricted quantifier (restricted to F, that is) which combines with a formula G to form a quantified formula [$Q_x : F$]G in which all occurrences of x are bound. The satisfaction conditions for such formulas with respect to sequences are analogous for the different quantifiers. For example, a sequence s satisfies [$the_x : F$]G iff the sequence satisfying F and differing from s at most at "x" also satisfies G (or, on a Russellian view, iff there is exactly one sequence satisfying F and differing from s at most at "x" which also satisfies G). A sequence s satisfies [$many_x : F$]G iff many sequences satisfying F and differing from s at most at "x" also satisfy G. And so on.

insane. Jones' utterances of (9a), (9b), (10a) and (10b) contain referential uses of "some murderers of Smith", "some guys in the dock", "a few murderers of Smith" and "a few guys in the dock", respectively.

However, I think it is clear that these referential uses of the quantifier phrases in (5)–(10) are not of an infrequent kind; or at any rate, they are not of any relevant kind less frequent than any relevant kind to which the referential use of "the murderer of Smith" in the relevant utterance of (1) above belongs. Nor are they abnormal or nonstandard; or at any rate, they are not abnormal or nonstandard in any sense in which the referential use of "the murderer of Smith" in the relevant utterance of (1) above would not be similarly abnormal or nonstandard. If these impressions are correct—and even if (to my surprise) they weren't—there are good reasons why they should be. For those utterances do not seem to be able to communicate the intended contents only in virtue of their being embedded in special contextual setups or in virtue of the audience's ability to navigate inferentially from a quirky utterance on the part of the speaker; or, at any rate, they do not seem to be able to communicate the intended contents in virtue of any feature of context and audience differing substantively from corresponding features of the context and audience in the case of the referential use of "the murderer of Smith" in the relevant utterance of (1) above. I think these claims are evident even already, but let me develop them briefly.

Arguably, there is nothing special about the contextual setups of the relevant utterances of (5)–(10). In particular, there is no semi-occult piece of information that must be common knowledge between Jones and her audience in order for the communication of suitable contents involving particular individuals to occur successfully. There are certainly many bits of information that must be common knowledge, as in any successful communicative event, but these would not seem to be substantively different from the fairly obvious bits of information that are required to be common knowledge in the case of the relevant utterance of (1): faced with an utterance of one of the sentences of the form [Q_x: *x is a murderer of Smith*] *x is insane* above, both Jones and her audience must know or be able to easily work out that the particular persons involved in the relevant communicated content are or provide, or are believed to be or provide, [Q_x: *x is a murderer of Smith*], just as in the earlier case both Jones and her audience must know or be able to easily work out that the person in the dock (Peters) is, or is believed to be, the unique murderer; and of course in all cases both Jones and her audience must assume that moving frantically in such a situation is a probable sign of insanity. But hardly more than this would seem to be required as common knowledge.

Presumably in part because of the rather non-special nature of the contextual setups of the imagined utterances of (5)–(10), a typical interlocutor of Jones will not need to exercise any inferential ability in order to grasp the contents involving particular persons that Jones intends to communicate; or, at any rate, such an interlocutor will not need to exercise any kind of inferential ability in the case of these utterances that he did not exercise in the case of the relevant utterance of (1). For example, as he hears Jones' utterance of (5a) and sees the people in the dock moving frantically, he will non-inferentially grasp that Jones intends to communicate to him that those people are insane, just as in the case of an utterance of (1) in the setup above he will non-inferentially grasp that Jones intends to communicate to him that Peters is insane. There would seem to be little if any role for inference also in the grasp by such an interlocutor of the intended contents of the imagined utterances of (5b) and (6)–(10).

The non-special character of the contextual setups for the imagined utterances of (5)–(10) and the ready accessibility of the communicated content to a typical interlocutor in the setups surely play a part in the impression that they are perfectly normal, standard utterances, lacking the air of quirkiness of the imagined utterances of (3) and (4). (At any rate, as noted above, it's hard to see in what way they would be less normal or standard than the relevant utterance of (1).) Presumably these characteristics of the setups and of what is required of the audience explain also to a large extent the impression that the imagined utterances of (5)–(10) are more frequent in kind than utterances relevantly similar to the imagined utterances of (3) and (4). (But perhaps only a large empirical study could fully confirm that cases of referential use involving special setups and requiring the exercise of moderately sophisticated inferential abilities on the part of the audience are in fact less frequent than uses such as those in the imagined utterances of (5)–(10), much as the former uses have an intuitive air of abnormality.)

It is worth noting that referential uses of descriptions may be of the presumably infrequent, nonstandard kind. Let's go back to the first Neale scenario and suppose Jones had instead uttered

3'. Well, the people taking my seminar turned up.

Surely with this utterance Jones could have fully intended to communicate to Neale that Smith (and only Smith) attended the party, but the utterance surely sounds odd given the setup. In fact, in this setup communication of the content involving Smith, while surely feasible, clearly requires the performance of an inference, however modest, on Neale's part—one allowing him to grasp that what Jones means is that Smith (and only Smith) attended the party. Or suppose Soames and Neale are having a conversation in Princeton in the late 1980s and Soames wants to know who among his acquaintances has played cricket; suppose further that at the time of Soames's and Neale's conversation Lewis is the only Princeton philosopher lecturing abroad, and this is vaguely remembered by Princeton philosophers; imagine Neale utters

4'. The Princeton philosopher who is now lecturing abroad has played cricket,

fully intending to communicate to Soames that Lewis has played cricket before. The utterance sounds pretty quirky again, and communicative success, while surely feasible, clearly requires the performance of an inference, however modest, on Soames's part—one allowing him to grasp that what Neale means is that Lewis has played cricket.

The claim that referential uses of quantifier phrases other than descriptions are frequent, can be perfectly standard and don't require special scenarios, can be further bolstered with the help of many other examples. Let me give another three batches, just to give a sense of how easy it is to generate such examples.

Let's first move to the other favorite Donnellanian scenario, a party where champagne and cocktails are offered. In an example of Kripke's inspired by a similar one of Donnellan's, we are to suppose that one of the men in a bunch of people at the party is drinking champagne,[5] and he looks happy; then a watcher can intend to communicate to a companion that *that man* is happy by uttering *The man drinking champagne over there is happy*, and easily do so. Let's consider the following variations on Kripke's example:

11. Every man drinking champagne over there is happy.
12. Most men drinking champagne over there are happy.
13. Many men drinking champagne over there are happy.
14. Several men drinking champagne over there are happy.
15. Some men drinking champagne over there are happy.
16. A few men drinking champagne over there are happy.

Now think of the watcher and a companion of hers looking amusedly at a small bunch of men, *a*, *b*, *c*, *d* and *e*, all drinking champagne. First, imagine that all of the champagne drinking men give evident signs of happiness, and the watcher then makes an utterance of (11). The watcher may intend to communicate, and will at any rate successfully manage to communicate to her interlocutor, that *a*, *b*, *c*, *d* and *e* are happy. This utterance contains a referential use of "every man drinking champagne over there". Next imagine that it's *a*, *b*, *c* and *d* that look happy, but *e* conspicuously doesn't, and the watcher makes an utterance of either (12), (13) or (14) intending to communicate, and successfully managing to communicate to her interlocutor, that *a*, *b*, *c* and *d* are happy. These utterances contain referential uses of "most men drinking champagne over there", "many men drinking champagne over there" and "several men drinking champagne over there", respectively. Finally, imagine that it's only *a* and *b* that look conspicuously happy, and the watcher makes an utterance of either (15) or (16) intending to communicate, and successfully managing to communicate to her interlocutor, that *a* and *b* are happy. These utterances contain referential uses of "some men drinking champagne over there" and "a few men drinking champagne over there", respectively. Again I think it will seem fairly clear that the setup doesn't require special common assumptions or the performance of even moderately sophisticated inferences on the part of the audience, nor do the utterances have any air of nonstandardness or of being of an infrequent kind.

Let's now turn to an even more undistinguished context than the ones considered so far, the local supermarket. Imagine that my wife and I are have just opened for inspection a carton of half-a-dozen eggs we intend to buy, and consider these sentences:

17. Every egg is broken.
18. Most eggs are broken.
19. Many eggs are broken.
20. Several eggs are broken.

[5] As recalled above, Donnellan's example discussion emphasizes the possibility, now widely thought to be irrelevant, that the man is not really drinking what the watchers think he is drinking (a martini, in Donnellan's example).

21. Some eggs are broken.
22. A few eggs are broken.

Suppose first that every egg in the carton we are looking at is evidently broken. If I then make an utterance of (17) intending to communicate to my wife that the eggs e1, e2, e3, e4, e5 and e6 (as we may call them) are broken, I will surely be successful in doing so. This utterance contains a referential use of “every egg”. Suppose now that e1, e2, e3 and e4 are evidently broken. My wife may make an utterance of either (18), (19) or (20) intending to communicate, and successfully managing to communicate to me, that e1, e2, e3 and e4 are broken. These utterances contain referential uses of “most eggs”, “many eggs” and “several eggs”, respectively. Finally, if only e1 and e2 are broken, I may make an utterance of either (21) or (22) intending to communicate, and successfully managing to communicate to my wife, that e1 and e2 are broken. These utterances contain referential uses of “some eggs” and “a few eggs”, respectively. Once more it seems clear that the setup is not particularly special or requires the exercise of even moderately sophisticated abilities on the part of the interlocutors, and the utterances do not have any appearance of nonstandardness or of being infrequent in kind.[6]

Here is a final bunch of examples involving a non-special setup and lacking an appearance of nonstandardness or infrequency in kind. Consider the following sentences:

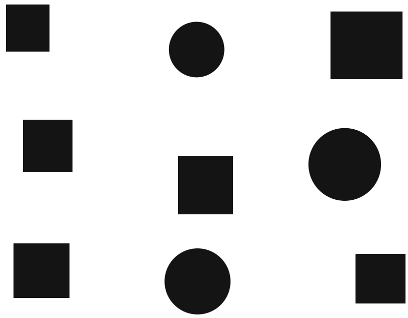
23. Every figure is black.
24. Most figures are squares.
25. Many figures are squares.
26. Several figures are circles.
27. Some figures are circles.
28. A few figures are circles.

Now think of two people contemplating Fig. 6.1 as they calmly describe what they see. (As in the vast majority of Donnellan’s original examples, we need not suppose that the interlocutors are engaging in anything more ambitious than calling each other’s attention to aspects of what they see.) Call the figures, from left to right and from top to bottom, A, B, C, D, E, F, G, H and I.

One of the interlocutors may make an utterance of (23) intending to communicate, and successfully managing to communicate to his interlocutor, that *A, B, C, D, E, F, G, H* and *I* are black. This utterance contains a referential use of “every figure”. One of the interlocutors may make an utterance of either (24) or (25) intending to

[6]Given that several authors already take indefinite descriptions to have frequent and standard referential uses, I have not emphasized this fact again with new examples. But note, e.g., that in the murderers’ setup, if only Adams is moving frantically, Jones can utter *A murderer of Smith is insane* and successfully communicate that Adams is insane; that in the party setup, if only *a* looks happy, the watcher can utter *A man drinking champagne over there is happy* and successfully communicate that *a* is happy; and that in the supermarket setup, if only e1 is broken, I can utter *An egg is broken* and successfully communicate that e1 is broken. All these are perfectly standard referential uses.

Fig. 6.1 A bunch of figures



communicate, and successfully managing to communicate to his interlocutor, that *A, C, D, E, G* and *I* are squares. These utterances contain referential uses of "most figures" and "many figures", respectively. Finally, one of the interlocutors may make an utterance of either (26), (27) or (28) intending to communicate, and successfully managing to communicate to his interlocutor, that *B, F* and *H* are circles. These utterances contain referential uses of "several figures", "some figures", and "a few figures", respectively.

It is tempting to conjecture that a sufficient condition for a referential use of a quantifier phrase $[Q_x : Fx]$ in an utterance of a sentence $[Q_x : Fx]Gx$ to be of the frequent, standard kind is that the particular object or objects about which the speaker intends to communicate that they are *G*s and of which the speaker means that they are or provide $[Q_x : Fx]$ are objects in plain view of speaker and audience, that the fact that those objects are or provide $[Q_x : Fx]$ is non-inferentially clear to speaker and audience, and that the fact that they are Gs is non-inferentially clear to speaker and audience. The question of what might be jointly necessary and sufficient conditions for a referential use of a quantifier phrase to be standard and of the frequent kind is bound to be harder to ascertain.

6.2 Constraining Theories of Referential Uses

There are three broad kinds of theories of referential uses of descriptions, which we might call *ambiguity* theories, *contextualist* theories and *pragmatic* theories. Ambiguity and contextualist theories postulate that the contents involving particular objects which are characteristically communicated in referential uses are literally or semantically expressed by the relevant utterances. Pragmatic theories postulate that these contents are not semantically expressed in these uses (or at least in the vast majority of them[7]), and are conveyed via some pragmatic mechanism from the literally expressed contents, which do not involve particular objects in the

[7]On the reason for this parenthetical qualification see Sect. 6.2.2 below.

relevant way. The common view that referential uses of quantifier phrases other than descriptions are different from referential uses of descriptions in the way recalled in Sect. 6.1 has recently given much impetus to ambiguity and contextualist theories of referential uses of descriptions. One reason, discussed in Sect. 6.2.1 below, is that, while the semantics of other quantifier phrases has generally always been thought to be strictly non-referential, the phenomena of standardness and frequency of their referential uses suggest that definite (and possibly indefinite) descriptions might exemplify different semantic conventions. A second reason, discussed in Sect. 6.2.2, is that the alleged differences between the referential uses of descriptions and those of other quantifier phrases suggest that a pragmatic explanation of referential uses of descriptions is not possible, even if it is feasible or even compulsory for referential uses of other quantifier phrases. In this section I will question the adequacy of these reasons, and suggest that the situation is less clear-cut than has lately been thought. Much of this discussion appeals to the data from Sect. 6.1 in order to state a number of constraints on the extensions of theories of descriptions to cover the referential uses of quantifier phrases in general.

6.2.1 Ambiguity and Contextualist Theories

Ambiguity theories of referential uses of definite descriptions postulate that the definite article “the” is lexically ambiguous as between an attributive meaning and a referential meaning (*mutatis mutandis* for indefinite descriptions, here and in what follows). When a speaker uses a definite description, she typically uses it either with an intention to use it attributively or with an intention to use it referentially, and the audience must disambiguate occurrences of “the” as a part of the process of communication involving utterances containing it, just as one must disambiguate occurrences of “bank” in communicative utterances containing it. At least for a good many years, ambiguity theories of definite descriptions were often thought to face a big problem in the existence of referential uses of other quantifier phrases of the kind illustrated by the examples involving (3) and (4): if the phenomenon of referential use takes place across the whole range of quantifier phrases, then it’s reasonable to expect that a common explanation for all quantifiers would be needed; but the postulation of a whole range of ambiguities involving many quantifiers would then seem methodologically unattractive and in need of substantive additional motivation, at least if some full pragmatic explanation of the same common phenomenon is available. And for a good many years, it was often thought that such a pragmatic explanation was available (see Sect. 6.2.2 below).

However, in recent times the methodological problem of ambiguity theories has been thought by many to be only apparent. As noted in Sect. 6.1, referential uses of quantifier phrases other than descriptions have been thought by many recent authors to be instances of a phenomenon significantly different from the phenomenon of referential uses of descriptions. In particular, it has been thought that referential uses of quantifier phrases other than descriptions are nonstandard and of an infrequent kind. Referential uses of descriptions, being frequent and standard, would then seem to be more plausibly explainable as relying on, or perhaps just being part of the conformation of, a distinctive referential meaning for the definite article without counterparts in other quantifiers (except possibly the indefinite article). This view is embraced by Reimer [20] and Devitt [5, 6], and is considered the most likely account (in view of these authors' arguments) in the recent textbook by Barbara Abbott [1]. Furthermore, the observation that referential uses of descriptions are frequent and standard at any rate strongly suggests that these uses cannot have the pragmatic explanation that the infrequent nonstandard referential uses of other quantifier phrases is bound to have (see Sect. 6.2.2 below). So it now appears somewhat common to think that the former use is simply not amenable to a pragmatic account.[8]

But the data reviewed in Sect. 6.1 suggest that, as far as frequency and standardness are concerned, the phenomenon of referential uses of quantifier phrases other than descriptions is not significantly different from the phenomenon of referential uses of definite descriptions, after all: all usual quantifiers give rise both to referential uses of a standard, presumably frequent kind, and to referential uses of a nonstandard, presumably infrequent kind. Consequently I think that, at least provided that some common pragmatic explanation can be given of the referential use of quantifier phrases in general, the situation will in fact be what it was generally thought to be for many years. This does not imply by itself that, even if the proviso about the existence of a pragmatic explanation can be met, ambiguity theories must be wrong, or even that they face an insuperable difficulty. But it does imply the constraint that ambiguity theories of definite descriptions must be extended via the postulation of similar ambiguities across the whole range of usual quantifiers. This postulation would thus appear to face the same problems of methodological motivation that it was for many years thought to face, at least if some full pragmatic explanation of the same common phenomenon is indeed available.[9] And I myself

[8]In this connection both Devitt [5, 6] and Reimer [20] give another influential consideration, based on the existence of "dead metaphors". To use Reimer's example, the verb "to incense" originally meant and still means to perfume by burning incense, but now has an additional meaning, "to make angry", derived from the standardization of the "incendiary" metaphor. The initially metaphorical meaning can surely be pragmatically worked out by a person competent with the original meaning, even if he is unfamiliar with the fact that the metaphor is now dead and the metaphorical meaning is now a literal meaning. But this pragmatic derivability of the initially metaphorical meaning surely does not imply that it is still purely metaphorical. Similarly, Devitt and Reimer claim, even if someone can pragmatically work out a referential use of a description from its attributive meaning, this would not mean that the referential use is not literal. According to Devitt and Reimer, this places a heavy burden on pragmatic explanations of referential use: they must show that pragmatic explanations are not merely possible, but somehow necessary—a burden that according to Devitt and Reimer has not been lifted. See a later note for a remark on these claims.

[9]There are several other somewhat indirect considerations against ambiguity theories. For example, as emphasized, e.g., by Kripke [12] and Bach [2], if the definite article were ambiguous, one would expect the ambiguity to be reflected in the existence in other languages of pairs of different words corresponding to its two meanings in English, but there is at best only flimsy evidence for this.

think that there is in principle no obstacle to such a pragmatic explanation, as I will explain below (in Sect. 6.2.2).

Some slightly more specific worries suggested by the similarities across the whole range of quantifiers affect particular ambiguity theories. Perhaps the best recent ambiguity theory is Devitt's [5, 6]. On this theory, the definite article has a meaning on which it works essentially as a demonstrative, and definite descriptions used referentially are essentially complex demonstratives. A description *the F* used with its referential meaning has an associated conventional rule that fixes the reference of an utterance of it in a context as the unique F that the utterance is "causally grounded in by perception" ([5], 292)[10]; then (in felicitous uses) the object itself comes to be involved in the literally expressed content, but the quantificational aspect that appears at the reference-fixing level does not make it into the literally expressed content. (By contrast, if *the F* is used attributively, the expressed content is quantificational, and does not involve a particular object in the relevant way.) One worry is then that in order to account for the presumably frequent and perfectly standard use of phrases of the form *every F* or *most F*s to communicate contents about particular objects (meant to provide every F or most Fs), such a theory ought to postulate meanings for "every" and "most" on which phrases of the form *every F* or *most F*s work essentially as demonstratives, with suitable associated conventional rules fixing the reference in context of *every F* and *most F*s. However, such a proposal, even if it turned out to be ultimately correct, would at this point appear to lack an independent motivation coming from semantic studies of the quantifiers.[11]

Contextualist theories postulate that the definite article is not ambiguous, having the same abstract meaning in all its uses. However, they postulate that this meaning must be supplemented by some concrete aspect of the context of utterance (in a very broad understanding of "context"), as a function of which the utterance of a definite description *the F* will literally express either a content involving some

[10]Recall that definite descriptions have referential uses of the infrequent, nonstandard kind, as pointed out in the discussion of examples (3') and (4'). Note that in these examples a definite description is being used referentially, and yet the audience can only work this out not by grasping how the utterance hints at some object that it (the utterance) is "causally grounded in by perception", but via a process apparently similar to that for working out particularized conversational implicatures of the "exploitative" type (see Sect. 6.2.2 below). The definite description is certainly not used here as a demonstrative. Thus it is not the case that referential uses of descriptions group together into the frequent and standard kind, and therefore if the definite article is to be held to be ambiguous as between a referential and a non-referential meaning, it must be held also that not all referential uses of definite descriptions fall under the referential meaning of the definite article. This seems a bit odd, for one might have expected to be enough for a speaker to use the definite article with its referential meaning that she should use it with the intention to use it referentially. Theories like Devitt's must hold that use of the article with its referential meaning requires more than a referential intention.

[11]By contrast, I am sympathetic to the idea that some quantifiers, such as "most", "many" and "all" (but not "every" or "the"), have bare pronominal uses with many characteristics of demonstratives, including aspects of reference fixing and modal profile. See Gómez-Torrente [9] for some brief remarks on this question.

particular object or objects (that the speaker means to be the F) or a general content not involving particular objects (meant by the speaker to be the F). These theories differ substantively from ambiguity theories in that they always postulate that an utterance of a definite description *the F* must express a distinctively quantificational component which is (part of) its minimal meaning; in non-referential uses this quantificational meaning is more or less left to stand alone and in referential uses it is somehow supplemented by a relevant object or objects (meant to be the F) or at least some "*de re* mode of presentation" of such object or objects.

The aspects as a function of which the utterance of a definite description will literally express either a general content or a content also involving some particular object or objects (meant to satisfy the description), as well as the details of the mechanism by which the supplementation proceeds in each case, vary somewhat with the different theories of this kind. In François Recanati [18, 19] the relevant aspect appears to be simply an intention on the part of the utterer, namely the intention whether to use the relevant description referentially or attributively. In Anne Bezuidenhout [3] and George Powell [17], the relevant aspect seems to be a more complex cluster of things, which includes speaker intentions that the audience grasp the relevance to utterance interpretation of some favored "mode of presentation" or "individual concept"; this may be quite general and inspecific (in attributive uses) or quite particular and specific (in referential uses).

A typical complaint against these contextualist theories is that they don't make it quite clear how a content involving particular objects is generated from the basic quantificational meaning in referential uses. The mystery is compounded when one notes that in all uses there must be at least a nucleus of quantificational meaning involved and literally expressed; how is this possible in referential cases, where the literally expressed content is supposed to involve some particular object or objects? (The typical ambiguity theorist doesn't have this problem, for he is free to postulate that the contents expressed in attributive uses and those expressed in referential uses are of thoroughly different kinds, the latter ones not needing to involve any quantificational component.)

Neale ([16], 171ff.; see also Ludlow and Neale [14], 303) proposes a contextualist theory designed in part to avoid this complaint (similar theories are embraced or viewed with sympathy by Paul Elbourne [8] and John Hawthorne and David Manley [11]). The idea is that the content literally expressed depends on the presence (in referential uses) or the absence (in non-referential uses) of an intended, largely non-descriptive, referential *completion* that effects a domain restriction of the description's matrix to a single individual (or designated group of individuals, in plural uses). Thus, for example, in the relevant referential use of (1) above, Neale says that a formal representation of the content literally expressed would look like this:

$$[the_x : x \textit{ is a murderer of Smith } \& \; x = a] \; x \textit{ is insane},$$

where a is some kind of representation that refers in some fairly non-descriptive way to the particular person in the dock. The completion "& $x = a$" would once more be provided by the speaker's intentions.

One quite general source of methodological dissatisfaction with contextualist theories of descriptions, analogous to the methodological source of dissatisfaction with ambiguity theories, is that contextualist theories must postulate fairly more sophisticated meanings or “semantic potentials” for the definite article than pragmatic theories. (In ambiguity theories, the dissatisfaction is that they postulate *more* meanings than pragmatic theories.) On a pragmatic theory, the definite article merely has the potential to generate a semantically expressed content of the quantificational kind (possibly involving some standard determinant of quantifier domain restriction); on a contextualist theory, the definite article must have a separate potential to generate a semantically expressed content of the kind that involves particular objects in the desired way (and if it does this via domain restriction, it must be via a rather special intended restriction to a single object[12]). However, for the same reasons as in the case of ambiguity theories, in recent times this methodological dissatisfaction has been thought by many to be only apparent. If referential uses of quantifier phrases other than descriptions are instances of a phenomenon significantly different from the phenomenon of referential use of definite descriptions, then it would not appear to be so objectionable methodologically to propose that the meaning or “semantic potential” of the definite article should be in a sophisticated class of its own (or perhaps one including only the indefinite article also). If referential uses of quantifier phrases other than descriptions are nonstandard and of an infrequent kind, the frequent and standard referential uses of descriptions are more plausibly explainable as relying on a distinctive referential “semantic potential” of the definite article without counterparts in other quantifiers (except perhaps the indefinite article). But once again, since the data reviewed in Sect. 6.1 are suggestive of fairly strict similarities between definite descriptions and other quantifier phrases, I think the methodological worry stands. (Again, modulo the supposition, vindicated in Sect. 6.2.2 below, that some fully pragmatic explanation of the same common phenomenon is indeed available.)

The contextualist literature barely considers the general quantifier case (unlike the initial polemical literature between ambiguity theorists and pragmatic theorists). But given that referential uses of quantifier phrases in general appear to be essentially analogous to referential uses of descriptions, a reasonable constraint on contextualist proposals about descriptions must be that they are to be extended to

[12]An intended restriction to a single object does seem rather special particularly in cases (of referential use of the frequent, standard kind) where the speaker’s intended domain restriction is explicitly not one to a single object. Thus, if in the party scenario one of the champagne-drinking men is wearing a pink suit and looks conspicuously drunk, and the watcher makes an utterance of

The man wearing a pink suit among those over there is drunk,

it would seem that it is unmotivated to postulate that the speaker somehow intends a further restriction via the property of being identical to that man, let alone one that the meaning of the definite article somehow forces her to provide given her intentions. Thus in such cases there are even clearer problems of motivation for postulating that the intuitively conveyed content involving that man (as wearing a pink suit and being drunk) is literally expressed by the watcher’s utterance.

analogous accounts of referential uses of quantifier phrases in general. This suggests a more substantive worry for what seem to be the best kind of proposals of this type, those in the spirit of Neale's suggestion. The worry is that, while the device of domain-restricting referential completions of the quantifier matrix appears to yield the intuitively correct referential truth conditions in the case of utterances containing definite descriptions, it is dubious that it can do so for all quantifiers susceptible of (frequent and standard) referential uses.

Thus, consider the relevant referential utterance of (6a). The matrix completions that immediately suggest themselves,

$[most_x : x \textit{ is a murderer of Smith} \ \& \ (x = \textit{Adams} \vee x = \textit{Barnes} \vee x = \textit{Crane} \vee x = \textit{Daniels} \vee x = \textit{Evans} \vee x = \textit{Foster} \vee x = \textit{Green})] \ x \textit{ is insane}$

and

$[most_x : x \textit{ is a murderer of Smith} \ \& \ (x = \textit{Adams} \vee x = \textit{Barnes} \vee x = \textit{Crane} \vee x = \textit{Daniels} \vee x = \textit{Evans})] \ x \textit{ is insane,}$

don't have the intuitively desired truth conditions; neither expresses a condition sufficient for Adams, Barnes, Crane, Daniels *and* Evans to be insane. The examples (7)–(10) give rise to similar problems.[13] In general, in many of these examples it just does not appear to be the case that the communication of the relevant content involving particular objects in the desired way is a matter of domain restriction, let alone one of "referential" domain restriction. Of course, the content that the particular individuals in which we are interested in each case are insane can be added outside the quantifier matrix. But resorting to this would appear to deprive the proposal for quantifiers in general of the motivation it has for the definite article, which is to indicate how the quantificational, "general" aspect of the meaning of the article interacts in a systematic, non-vacuous way with the "particular" aspect of content in the truth conditions of referentially used utterances of sentences containing descriptions. Surely contents involving particular objects must appear somehow in the theoretical picture, but to postulate that they just pop up without significant interaction with the quantifiers ought to seem non-explanatory from within the contextualist frame of mind.

[13]By contrast, "every" seems amenable to the completion treatment:

$[every_x : x \textit{ is a murderer of Smith} \ \& \ (x = \textit{Adams} \vee x = \textit{Barnes} \vee x = \textit{Crane} \vee x = \textit{Daniels} \vee x = \textit{Evans} \vee x = \textit{Foster} \vee x = \textit{Green})] \ x \textit{ is insane}$

seems to have the intuitive truth conditions of the intended content communicated by a referential use of (5a).

6.2.2 Pragmatic Theories

Pragmatic theories of the referential uses of definite descriptions propose that these are not ambiguous and are not sensitive to a contextual aspect as a function of which they come to literally express particular or general contents. (But they are compatible with the possibility that they may be context-sensitive in some more standard way, e.g. that they may be sensitive to some contextual determinant restricting the domain of quantification.) Instead, pragmatic theories propose that an utterance containing a definite description used referentially will in the vast majority of cases literally express a content that is not the communicated content involving in the desired way the relevant particular object(s), but that the audience typically can and do grasp this content via some pragmatic mechanism. Note that this characterization is deliberately compatible both with the possibility that on a pragmatic theory the content of an utterance containing a referentially used description may be a not purely general content (even a content that somehow involves the object denoted by the description), and with the possibility that a pragmatic theory may accept that in some (presumably rare) cases the communicated content involving a particular object is the same as the semantically expressed content. As for the first possibility, simply recall our example from Sect. 6.1 involving the description "the man among those over there who will get here first". As for the second possibility, think of a similar example, in which there is in fact just one man in plain view that the speaker is looking at, and she now makes an (arch) utterance of

The man identical to that man will get 10 dollars,

It is open to the pragmatic theorist, as understood here, to accept that a content involving in the desired way that particular man, which is presumably communicated by this utterance, is also semantically expressed by it.[14]

The data in Sect. 6.1 again suggest that pragmatic theories must work under the constraint that referential uses of descriptions and of other quantifier phrases ought to be explained by a similar mechanism or mechanisms. The data make it unlikely that substantively different mechanisms should be involved in accounting for referential uses of quantifier phrases containing different quantifiers. But this poses a considerable problem to pragmatic theories, as these have typically covered only what we have called the infrequent, nonstandard referential uses of quantifier phrases (including descriptions). In fact, some original proponents

---

[14]Similarly, a pragmatic theory of quantifier phrases in general may accept that in some (presumably rare) cases the communicated content involving some particular object(s) is the same as the content semantically expressed by an utterance of a sentence containing a quantifier. Think again of the three men example, and of an (arch) utterance by the speaker of

Every man identical either to this man, that man or that other man will get 10 dollars.

A pragmatic theorist as understood here can accept that a content involving in the desired way those particular men, which is presumably communicated by this utterance, is also semantically expressed by it.

of pragmatic theories, like Neale ([16], 171ff.; see also Ludlow and Neale [14], 303), have deserted the pragmatic camp apparently to a great extent because only the infrequent, nonstandard referential uses of quantifier phrases in general seem susceptible of receiving a straightforward pragmatic explanation.[15]

Let's begin by recalling what this straightforward pragmatic explanation is. It involves the mechanism of communication of particularized implicatures via a "flouting" of the maxims of conversation, the mechanism Grice called "exploitation". In such cases, as we may recall, an audience might in principle reason in this way when facing a quirky utterance with the literal content that *p*—even if as a matter of fact no one reasons in this way outside the confines of a philosophy classroom or a philosopher's study: "In saying that *p* the speaker appears to be in some affected way infringing the plausible maxims regulatory of conversation identified by Grice; however, we presume that she is nevertheless attempting to follow the maxims; assuming that she is not really seeking to infringe the maxims, we must suppose that she is attempting to communicate a certain content *q* different from *p*, and to be doing so precisely because she must know that we can see or infer this". Neale ([15], 88–9) suggests an explanation of this form for the cases that concern us, and in fact for all referential uses of quantifier phrases, when he gives a fleshed out Gricean derivation for a case similar in all relevant respects to the ones involving (3') and (4') above, saying also that this is the same kind of explanation required by the first cricket example, the one involving (4) above. Specifically, Neale suggests that the crucial sub-reasoning in which the hypothesis that the speaker attempts to communicate a certain content *q* involving a particular object is worked out will appeal to the Gricean maxims of Relation and of Quality; more specifically

[15]Devitt [5, 6] has argued on very general grounds that no pragmatic explanation can work. His idea is that in cases where *the F* is incomplete and the speaker has "beliefs about x [the object the speaker has in mind] that are too inadequate—ignorance—or too wrong—error—to enable her to supply the completion demanded by the explicit approach [to quantifier domain restriction] or to delimit the domain of quantification as demanded by the implicit approach" ([6], 18), she will still successfully use *the F* to convey a thought about x. Devitt continues:

> Where 'the F' is obviously incomplete, as it very often is, the speaker clearly does not believe or mean to say that there is an object that is uniquely F and is G; for example, that there is one and only one table in the world and it is covered with books. And the arguments from ignorance and error show that she is often not in a position to modify that general proposition, by completing 'the F' or delimiting its domain, into one that she might plausibly believe and mean. ([6], 19)

I find this unpersuasive. In referential uses speakers will typically be able to restrict the domain of quantification to, e.g., "table here", "eggs among these" or "figures over there"; Devitt's use of "often" in the next-to-last sentence quoted is baffling. It is especially baffling given Devitt's later claim that "often the speaker could have used a name, a simple demonstrative, or pronoun. So, when a definite is used referentially, there are nearly always other devices available" ([6], 20); surely the same devices could be used to effect suitable domain restrictions. Also, I fail to see why, in those presumably rare cases in which the speaker is wrong that the restriction she is able to provide helps pick the right object, it is not nevertheless the case that the speaker believes and means a general proposition; surely this is no problem for a pragmatic account.

still, Quality is crucially involved in that, according to Neale, when reasoning that the speaker must have adequate grounds for what she attempts to communicate, the (ideal, reasoning) audience will realize that she cannot plausibly be supposed to have purely general grounds for the general content *p*, and therefore must have the particular grounds *q*.

Here is Neale's fleshed out derivation:

(a) *S* [the speaker] has expressed the proposition that [the $x : Fx$](Gx).
(b) There is no reason to suppose that *S* is not observing the CP [Grice's Cooperative Principle] and maxims.
(c) *S* could not be doing this unless he thought that *Gb* (where '*b*' is a name). Gloss: On the assumption that *S* is observing the Maxim of Relation, he must be attempting to convey something beyond the general proposition that whoever is uniquely *F* is *G*. On the assumption that *S* is adhering to the Maxim of Quality, he must have adequate evidence for thinking that the *F* is *G*. I know *S* knows that *b* is the *F*, therefore *S* thinks that *Gb*.
(d) *S* knows (and knows that I know that he knows) that I know that *b* is the *F*, that I know that *S* knows that *b* is the *F*, and that I can see that *S* thinks the supposition that he thinks that *Gb* is required.
(e) *S* has done nothing to stop me thinking that *Gb*.
(f) *S* intends me to think, or is at least willing to allow me to think, that *Gb*.
(g) And so, *S* has implicated that *Gb*. ([15], 89)

Of course, Gricean derivations are never carried out by anyone outside the confines of a philosophy classroom or a philosopher's study, but it is nevertheless the case, as noted in Sect. 6.1 above, that in the setup of the relevant utterance of (4'), communication of the content involving Lewis requires Soames to perform a modest inference allowing him to grasp that what Neale means is that Lewis has played cricket before. This is presumably due to the fact that Neale makes an intuitively quirky utterance and sets Soames's mind working to reach (easily) the idea that he after all means that Lewis has played cricket before. So there is a connection between the fact that what Grice's theory requires in these cases is a certain kind of derivation and the fact that actual audiences in these cases will perform inferences and will not immediately detect the communicated content. Grice gives an explanation of the quirkiness with his idea that, as he puts it, the speaker "flouts" (or appears to flout) the implicit Gricean conversational maxims; and in cases of maxim flouting, cases of "exploitative" particularized implicatures, not only will a certain kind of reasoning be required in ideally rational Gricean derivations of the implicated content, but a certain amount of inference will be required of a flesh-and-blood audience as well.

Although this looks like a correct explanation of the nonstandard referential uses in question, the authors who have called attention to the frequency and standardness of the referential uses of descriptions have made it clear that the uses they have in mind are not explainable in this way. In these latter uses it seems clear that there is no inference involved, presumably in part because there is no quirkiness that could set it in motion. However, those authors have gone on to see this as the basis for embracing ambiguity or contextualist theories. The data of Sect. 6.1 suggest that this reaction, even if it need not be ultimately wrong, is certainly premature.[16] However, the conclusion to be drawn from these data is definitely not that things are back where they stood for the pragmatic theorist. The reasonable upshot is that, even if nonstandard referential uses of quantifier phrases in general are explainable by the mechanism of communication of exploitative particularized implicatures, the standard uses of the frequent kind cannot. Given this strict constraint, are there any alternatives open to the pragmatic theorist?

The standard, frequent kind of referential uses of quantifier phrases do not seem explainable by the mechanism of communication of generalized conversational implicatures either. Generalized implicatures are supposed to arise *normally* when an utterance of a given sentence is made, to arise in a default way as a function of the sentence alone, failing to arise only when the context incorporates some special feature or other (Grice [10], 37). But in utterances of a sentence containing a quantifier phrase, the kind of content communicated in referential uses is not communicated as a norm. Surely, e.g., the intended (kind of) content communicated by the referential utterance of (1) is not communicated as a norm by utterances of (1). It does not arise in a default way by any means, and in fact it obviously is communicated only as a function of special details of the context in each case. (Recall the two kinds of context for the utterances of (1) above.)[17]

The data thus suggest that, if explainable within a Gricean framework at all, the standard, frequent kind of referential uses of quantifier phrases must be explained

[16] As noted above, this reaction has been bolstered by Devitt's and Reimer's consideration involving "dead metaphors", which suggests to them that pragmatic theories must take on the burden of showing that a pragmatic explanation of referential use is not merely possible but necessary. As we saw, they think this because in dead metaphors a pragmatic explanation is possible (at least for some speakers) but clearly not necessary because incorrect. But I don't think this is right. The consideration of dead metaphors merely shows that the pragmatic theorist must indicate *some respect* in which the case of referential uses of quantifier phrases is *relevantly different* from the case of "dead metaphors" meanings. And he surely can. For example, as noted by Bach ([2], 226), clearly many speakers are competent with the "dead metaphors" meaning of a word but not with its original meaning; however, it does not appear to be the case that any speaker who competently grasps referential uses of quantifier phrases fails to grasp their non-referential uses. Also, as recalled in an earlier note, while the dichotomy between attributive and referential uses of the same quantifier is thoroughly cross-linguistic, dead metaphors are hardly ever cross-linguistic. Such differences immediately license the consideration of general explanations of referential uses lacking the evidently specific character of the ambiguity explanations appropriate for dead metaphor cases. In fact, to some the differences may even immediately suggest the inadequacy of ambiguity explanations of referential uses.

[17] Bach ([2], 227) sees standard referential uses of descriptions as explainable by the mechanism of generalized conversational implicatures. Neale ([16], 173) says in passing that his account in [15] (of the nonstandard uses) is "a generalized conversational implicature story". As just noted, I don't think this would be the correct explanation of standard referential uses. But in fact, as noted above, it is quite clear that Neale's ([15]) account, which is designed for what with hindsight we can see as nonstandard uses, is, quite appropriately, an "(exploitative) particularized conversational implicature story". Besides the remarks above, note that his Gricean derivations exploit the maxim of Relation; I agree with Stephen Levinson ([13], 127) that it is clear that all implicatures that arise from observing this maxim must be particularized.

as cases in which particularized implicatures are communicated, though not via "exploitation" (the "flouting" of maxims). In the Gricean framework these cases may be of two kinds. The first kind are the particularized conversational implicatures that Grice ([10], 32) characterizes as "examples in which no maxim is violated, or at least in which it is not clear that any maxim is violated", and which Levinson ([13], 104) has called "standard" particularized implicatures; while Levinson's use is revelatory of one right sort of connection with the slightly theoretical use of "standard" that we have made in this paper, to avoid confusion let's call these implicatures "non-violatory". These are implicatures where the communicated content is an amplification of what has been said that will typically be evident to a typical audience given the details of the context. By contrast with exploitative particularized conversational implicatures, non-violatory implicatures will often be intuitively obvious to the (typical) audience, without need of the exercise of any inferential ability. And yet, they don't need to arise normally or by default when an utterance of the relevant sentence is made, but only due to the special features of the context. Grice's paradigm example, let's recall, is that of a conversation in which a stranded motorist tells a passer-by that he is out of gas, and she replies with an utterance of *There is a garage round the corner*. (Part of) the implicated content, that will be obvious to the stranded motorist, is that the garage is open or believed to be open. Note that an utterance by a concerned passer-by of *There is a garage round the corner* would intuitively carry the same non-violatory implicature even if it were not prompted by any utterance of the stranded motorist.

That non-violatory implicatures are particularized of course doesn't mean that groups of them cannot be included in sets all of whose members share important properties. For example, the amplification invited by the speaker may be of a common type instantiated in many other cases. It is reasonable to suppose that one such common type of amplification will be that of *evident grounds* for the literally expressed content. For example, think of a conversation in which two people are contemplating the countryside and as he looks at the dark clouds in the horizon, one of them makes an utterance of *It's going to rain a lot*, intending to communicate also that there are dark clouds in the horizon. He will easily manage to communicate this content, as it provides (part of) the evident grounds in the context for the literally expressed content. This is clearly a non-violatory implicature, for surely the speaker is violating no maxim when he says that it's going to rain a lot. Note also that it's clearly not a generalized implicature, as utterances of *It's going to rain a lot* do not carry it by default (not nowadays, anyway, when we have all sorts of long-term weather forecasts).

The second kind of particularized non-exploitative implicatures are the implicatures that Grice ([10], 32) characterizes as "examples in which a maxim is violated, but its violation is to be explained by the supposition of a clash with another maxim". Let's call these "violatory non-exploitative" implicatures. Grice's example is a conversation in which two friends are planning a trip to France, and one of them would like to take the opportunity to visit a friend who lives in that country provided she lives somewhere close to the itinerary they have planned; when he asks his traveling companion where the friend lives, the companion utters *(She lives)* *somewhere in the South of France*. Grice suggests that the companion violates the maxim of Quantity (though he doesn't "flout" it), and if he does so it's because he doesn't want to violate the maxim of Quality ("Don't say what you lack adequate evidence for"); (part of) the implicated content is that the companion doesn't know in which town the friend lives. Again, by contrast with exploitative implicatures, violatory non-exploitative implicatures may well be intuitively obvious to the (typical) audience, without need of the exercise of any inferential ability. And again they don't need to arise by default when an utterance of the relevant sentence is made. Furthermore, that violatory non-exploitative implicatures are particularized again doesn't mean that they cannot group into sets all of whose members share important properties.

Might the standard referential uses of quantifier phrases involve non-violatory implicatures or violatory non-exploitative implicatures? Take this last possibility first, and suppose that Jones when making a referential use of, say, (5a) or (5b), implicates that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane. Now it might perhaps be suggested that she may be interpreted as violating the maxim of Quantity which directs her to provide as much information as is required, and at the same time she may be supposed to do this so as not to violate the maxim of Manner which directs her to be brief, avoiding unnecessary prolixity. If this is correct, then the prolixity of saying that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane will indeed be unnecessary, as the implicated content will be evident anyway as an obvious amplification of what she says, an amplification constituted in fact by (part of) the evident particular grounds for what she says. But I don't think this would be a correct explanation, for it is unclear that in the context of Jones' utterances of (5a) and (5b) there is a pre-existing expectation that she should provide more information than she provides with her utterances (under the assumption that their literal content is in a relevant way less informative than a related content involving Adams, Barnes, Crane, Daniels, Evans, Foster and Green). So it is dubious that she can be charged with violating in any clear way the maxim of Quantity. Similar remarks apply to the relevant uses of (1) and (6)–(10), and indeed to the relevant uses of (11)–(28).

Whether an analysis in terms of violatory non-exploitative implicatures is possible or not, an analysis in terms of non-violatory implicatures seems more feasible. On this alternative, Jones, when uttering (5a) or (5b) in the setup above, would not be violating, or at any rate *clearly* violating, any maxim, and instead should be interpreted as attempting to communicate also that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane by uttering (5a) or (5b), just because she realizes that in the context this is an evident amplification of what she in fact says that the audience cannot fail to grasp. And indeed, the audience will intuitively grasp that Jones could not be uttering (5a) or (5b) unless she thought that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane, for this is (part of) the evident grounds for what she says. Note that the relevant uses of (1) and (6)–(10), and indeed the relevant uses of (11)–(28), appear to be amenable to this treatment without significant variations.[18] Note also that the implicatures in these cases, if they are indeed such, don't arise normally or by default when utterances of the relevant sentences are made, even if referential uses are standard and frequent in kind. Finally, note that if the relevant contents are indeed communicated as implicatures of this kind, these implicatures probably belong together in several sets all of whose members share important properties, e.g. the set of non-violatory implicatures where the implicated content constitutes in context (part of the) evident grounds for the literally expressed content, the set of non-violatory implicatures where the implicated content constitutes in context (part of the) evident grounds for a quantificational part of the literally expressed content, etc.[19]

Like any other conversational implicature, non-violatory implicatures must be able to be derived in the Gricean fashion, even if as a matter of fact no speaker conducts these derivations in the normal business of language use. Here is a possible Gricean derivation of the implicature of the relevant utterances of (5a) and (5b) as a non-violatory implicature:

We presume the speaker to be following the plausible maxims regulatory of conversation identified by Grice when she literally expresses the content that $[every_x : x \textit{ is a murderer of Smith}]$ *x is insane.*

But we must suppose that he is aware of or thinks that *Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane* if we are to hold on to this presumption. (For she must be observing the maxim of Quality directing her to say only things for which she has adequate evidence, and the fact that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane is (part of) the manifest evidence she has for what she said.)

[18] I say that the relevant content involving particular objects must be "part of" the evident grounds for the literally expressed content, because there must be other grounds as well. Specifically, in the case involving (5a), besides the belief that Adams, Barnes, etc. are insane there must be a belief (common to speaker and hearer) that Adams, Barnes, etc. are all the murderers of Smith; in the example with (1), besides the belief that Peters is insane there must be a belief (common to speaker and hearer) that Peters murdered Smith by himself; etc. (For a more general remark concerning this kind of grounds see the last paragraph of this section in the main text.) Note also that the pragmatic explanation considered in the text is not committed to the clearly false claim that the existence of evident grounds involving particular objects is *sufficient* for a use of a given quantifier phrase to be referential; surely there are attributive uses made on grounds involving particular objects.

[19] Another of the considerations recently developed by Devitt against pragmatic theories of the referential use of definite descriptions is based on the claim that such theories must presuppose that the non-referential use preceded in time the referential use. They must presuppose this, according to Devitt, because the convention establishing the possibility of non-referential use "features in pragmatic explanations of referential use" ([6], 19). However, argues Devitt, we have no evidence that non-referential uses preceded referential uses, and perhaps both uses arose together. I think that Devitt's consideration does no harm to an analysis in terms of non-violatory implicatures. Note that one doesn't need to suppose that the uses of *It's going to rain a lot* that don't carry the implicature of its evident grounds must in any substantive temporal sense precede their uses to communicate that it's cloudy. There didn't have to be a single use that didn't carry this implicature before there could be uses that did.

And surely she thinks that we can see this, so she thinks that we can see that she thinks that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane.
So she has conversationally implicated that Adams, Barnes, Crane, Daniels, Evans, Foster and Green are insane.

Mutatis mutandis for the conjectured non-violatory implicatures of the relevant utterances of (1) and (6)–(28).[20]

An explanation of this form is certainly a general explanation promising to encompass the frequent and standard kind of referential uses of usual quantifiers. More specifically, on close inspection the explanation can be seen to rely on a common aspect of the semantics of quantifiers, one which immediately creates a potentiality for a certain kind of contents to be pragmatically communicated in standard ways, in non-*recherché* contextual setups, through utterances of quantificational sentences. Recall that the standard satisfaction conditions for the (conjectured literally expressed content of) formulas of the form $[Q_x : F]\ G$ with respect to sequences are analogous for the different quantifiers: in general, a sequence s satisfies a formula $[Q_x : F]\ G$ iff Q sequences satisfying F and differing from s at most at "x" also satisfy G. In the cases of standard referential uses of a sentence of the form $[Q_x : Fx]\ Gx$ that we have considered, this amounts to the truth condition of $[Q_x : Fx]\ Gx$ being that some objects are or provide $[Q_x : Fx]$ which at the same time are Gs. In our cases, furthermore, it is non-inferentially clear to speaker and audience that certain particular objects are or provide $[Q_x : Fx]$ and also that those same objects are Gs. This constitutes essentially the evident grounds for the relevant utterance of $[Q_x : Fx]\ Gx$ in context. Thus, that the relevant particular objects are Gs is a part of the contextually evident grounds for the utterance, and consequently is a content that is easily available for communication through the utterance—under the implicit assumption that the speaker has adequate grounds for what she says and says what she says in part because of this. Clearly, this kind of conjectured implicated content is so closely related epistemically to the conjectured literally expressed quantificational content, that it is plausible to suppose that it is

[20]The explanation in terms of non-violatory implicatures has an additional "exegetical" virtue: it helps to make Gricean sense of some of Kripke's ([12], 262ff.) appeals to his distinction between "semantic reference" and "speaker's reference". Kripke clearly intends at least some of these appeals to fall under an explanation in terms of Gricean conversational implicatures, and indeed in terms of particularized conversational implicatures, but he gives no details. Take the famous case of Jones raking the leaves: "Two people see Smith in the distance and mistake him for Jones. They have a brief colloquy: 'What is Jones doing?' "Raking the leaves'." (Kripke [12], 263) When the second person utters *(Jones is) raking the leaves*, the literally expressed content is surely that Jones is raking the leaves, and yet he also intends to communicate the content that man over there is raking the leaves. This content is plausibly communicated as a non-violatory particularized conversational implicature in the context, and in fact one that constitutes part of the contextually evident grounds for the speaker's utterance (another part being her false belief that Jones is that man over there). (Kripke's main example to illustrate Grice's distinction between literal meaning and speaker's meaning is also clearly an example of a non-violatory implicature: a burglar says to another *The police are around the corner*, implicating that they must split, in this case an evident relevant consequence in the context of the literally expressed content.)

easily communicated in the form of non-violatory implicatures in the whole range of cases that have been concerning us. It is in this sense that the pragmatic explanation in terms of non-violatory implicatures sketched above appeals to a general aspect of the semantics of the quantifiers.

6.3 Concluding Remarks

Let me finish with some brief remarks on the question of whether or to what extent the preceding considerations favor any particular kind of theory of the referential use of quantifier phrases over the others. My calling attention to the feasibility of a certain pragmatic (indeed, classical Gricean) explanation of the frequent and standard referential uses of quantifier phrases may suggest that I embrace this explanation in particular, perhaps on methodological grounds. It is therefore worth stressing that while I suspect that this explanation is correct, I nevertheless see the question of what is the right account of referential uses as fairly open.

I do take the data assembled in Sect. 6.1 and the discussion in Sect. 6.2 to prove that the theoretical situation regarding referential uses of quantifier phrases in general is more complex than usually thought of late, and in particular to prove that there is no series of data or theoretical considerations in the recent literature I am aware of suggesting unequivocally the need for an ambiguity or contextualist theory. Thus, I take the data assembled in Sect. 6.1 and the discussion in Sect. 6.2 to suggest that the situation is not too different from what I take was the situation for many years: semantic theories of referential use (either ambiguity or contextualist) seem empirically feasible, but pragmatic theories seem empirically feasible as well, and essentially only methodological considerations of theoretical economy and the like would seem to favor clearly a pragmatic explanation.

While to some sympathizers of pragmatic theories this may seem enough to tip the balance in favor of a pragmatic account, to me it signals, at best, that there is more work to be done, or, at worst, that the data and our best theoretical considerations might well be compatible with a situation of theoretical indeterminacy. At best, there is more work to be done, for I don't think that considerations of economy should be taken as more than minimally reliable grounds for choosing one theory over another; if that's all we have to go by, we must keep trying to find direct evidence that may adjudicate the issue between semantic and pragmatic theories. At worst, the situation may be one in which the apparatuses of classical semantics and pragmatics are just too coarse-grained to extricate what's really going on in an illuminating way; if the data underdetermine all theories we can come up with using those apparatuses, that would be bad news for typical theorists of referential use, for it would suggest that a fully different apparatus is needed to account for it. Let's hope that time will throw light on the confusing aspects of our present situation.

Acknowledgements Research supported by the Mexican CONACyT (CCB 2011 166502) and by the Spanish MICINN and MINECO (research projects FFI2008-04263 and FFI2011-25626).

References

1. Abbott, B. 2010. *Reference*. New York: Oxford University Press.
2. Bach, K. 2004. Descriptions: Points of reference. In *Descriptions and beyond*, ed. M. Reimer and A. Bezuidenhout, 189–229. Oxford: Clarendon Press.
3. Bezuidenhout, A. 1997. Pragmatics, semantic underdetermination, and the referential/attributive distinction. *Mind* 106: 375–410.
4. Davies, M. 1981. *Meaning, quantification, necessity*. London: Routledge and Kegan Paul.
5. Devitt, M. 2004. The case for referential descriptions. In *Descriptions and beyond*, ed. M. Reimer and A. Bezuidenhout, 280–305. Oxford: Clarendon Press.
6. Devitt, M. 2007. Referential descriptions and conversational implicatures. *European Journal of Analytic Philosophy* 3: 7–32.
7. Donnellan, K. 1966. Reference and definite descriptions. *Philosophical Review* 75: 281–304.
8. Elbourne, P. 2005. *Situations and individuals*. Cambridge: MIT.
9. Gómez-Torrente, M. 2008. Are there model-theoretic logical truths that are not logically true? In: *New essays on Tarski and philosophy*, ed. D. Patterson, 340–368. Oxford/New York: Oxford University Press.
10. Grice, P. 1975. Logic and conversation. In *Syntax and semantics. Vol. 3: Speech Acts*, ed. P. Cole and J. Morgan, 41–58. New York: Academic. Reprinted in Grice, P. 1989. *Studies in the way of words*, 22–40. Cambridge: Harvard University Press. References to the reprint.
11. Hawthorne, J., and D. Manley. 2012. *The reference book*. Oxford: Oxford University Press.
12. Kripke, S. 1977. Speaker's reference and semantic reference. *Midwest Studies in Philosophy* 2: 255–276.
13. Levinson, S. 1983. *Pragmatics*. Cambridge/New York: Cambridge University Press.
14. Ludlow, P., and S. Neale. 2006. Descriptions. In *The Blackwell guide to the philosophy of language*, ed. M. Devitt and R. Hanley, 288–313. Oxford: Blackwell.
15. Neale, S. 1990. *Descriptions*. Cambridge: MIT.
16. Neale, S. 2004. This, that, and the other. In *Descriptions and beyond*, ed. M. Reimer and A. Bezuidenhout, 68–182. Oxford: Clarendon Press.
17. Powell, G. 2001. The referential-attributive distinction. A cognitive account. *Pragmatics and Cognition* 9: 69–98.
18. Recanati, F. 1989. Referential/attributive: A contextualist proposal. *Philosophical Studies* 56: 217–249.
19. Recanati, F. 1993. *Direct reference. From language to thought*. Oxford: Blackwell.
20. Reimer, M. 1998. Donnellan's distinction/Kripke's test. *Analysis* 58: 89–100.
21. Sainsbury, M. 1979. *Russell*. London: Routledge and Kegan Paul.

Chapter 7
Quantification and Logical Form

Andrea Iacona

Abstract This paper deals with the logical form of quantified sentences. Its purpose is to elucidate one plausible sense in which a considerably wide class of quantified sentences can be expressed in a classical first order language. Sections 7.1 and 7.2 provide some preliminary clarifications. Section 7.3 illustrates by means of familiar examples how the truth conditions of quantified sentences can formally be represented. Sections 7.4 and 7.5 show that the method of formalization suggested is consistent with some established undefinability results, and that it can easily be extended to a broad variety of cases. Section 7.6 draws a distinction between logical and non-logical quantifier expressions. Finally, Sect. 7.7 adds some concluding remarks.

7.1 Two Questions Instead of One

The line of thought that underlies this paper stems from the idea that there is a crucial ambiguity in the question of what is the logical form of quantified sentences. This question can be construed in at least two ways:

(Q1) How are quantified sentences to be formally represented in order to account for the logical relations involving them?

(Q2) How are quantified sentences to be formally represented in order to provide a compositional account of their meaning?

At least *prima facie*, (Q1) and (Q2) are independent questions: one thing is to provide a formal explanation of the logical relations involving certain sentences, quite another thing is to provide a compositional account of the meaning of those sentences. However, the most common attitude towards (Q1) and (Q2) is to think

A. Iacona (✉)
Centro di Logica, Linguaggio e Cognizione, Dipartimento di Filosofia e Scienze dell'Educazione, Università di Torino, Via S. Ottavio 20, Torino, Italy
e-mail: andrea.iacona@unito.it

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_7

that they are closely related, in that one and the same notion of logical form can provide an answer to both. As it will be explained in this section, the line of thought advanced here differs from two major views characterized by that attitude: one is old, the other is new.

According to the old view, which goes back to Frege, (Q1) is prior to (Q2), in that the notion of logical form that proves adequate to address (Q1) also provides an answer to (Q2). Consider the following sentences:

1. All philosophers are rich
2. Aristotle is rich

Frege suggested that there is a substantial difference between (1) and (2): although (1) is superficially similar to (2), its logical form substantially differs from that of (2). The difference that Frege had in mind turns out clear if (1) and (2) are formalized in a classical first order language. Let L be a first order language whose vocabulary includes a set of predicate letters $P, Q, R \ldots$, a set of individual constants $a, b, c \ldots$, a set of variables $x, y, z \ldots$ and the connectives $\sim, \supset, \vee, \wedge, \forall, \exists$. (1) and (2) can be represented in L as follows:

3. $\forall x(Px \supset Qx)$
4. Pa

Here P stands for 'philosopher' and Q stands for 'rich'. If one regards this formalization as a guide to a compositional account of the meaning of (1), one will be inclined to think that, once we have an answer to (Q1), we also get an answer to (Q2).[1]

However, some doubts might be raised in connection with this view. First of all, it is not clear how (3) can figure as part of a compositional account of the meaning of (1), given that it does not explain the apparent semantic analogy between (1) and (2). (1) contains a noun phrase, 'all philosophers', which in many respect resembles 'Aristotle', while it does not contain the expression 'if…then…'. Secondly, even if (3) were regarded as the real semantic structure of (1), in spite of such disanalogies, it would still be an open question how a compositional account of (3) could be given. As it is well known, a definition of truth for the sentences of L can be provided in the way suggested by Tarski, assuming that the truth value of any formula $\forall x\alpha$ depends on the satisfaction conditions of α. However, Tarski's method does not guarantees compositionality. Since $\forall x\alpha$ is formed by adding $\forall x$ to α in accordance with the usual syntactic rule, in order for compositionality to hold, the truth value of $\forall x\alpha$ should result from the combination of the meaning of $\forall x$ with the meaning of α. But if truth in L is defined in the way outlined by Tarski, it is quite natural to read $\forall x$ as an expression that does not have meaning in isolation.[2]

[1] This line of thought originates from Frege [4].

[2] Note that no clear alternative to this reading is provided by Frege's notion of "second-level function". One might be tempted to say that $\forall x$ denotes a second-level function F, so that the truth value of $\forall x\alpha$ is obtained by combining F with the meaning of α. But this is not a viable

According to the new view, which is currently adopted within formal approaches to natural language, (Q2) is prior to (Q1), in that the notion of logical form that proves adequate to address (Q2) also provides an answer to (Q1). In this case it is assumed that logical form is determined by syntactic structure, where syntactic structure is understood as LF, that is, as a formal representation that is distinct from surface structure and is the input of semantic interpretation. The LF of (1) and (2) may be represented as follows in order to provide a compositional account of their meaning:

5. [Every philosopher$_1$[t_1 is rich]]
6. [Aristotle[is rich]]

If one assumes that the logical form of a sentence is determined by its LF, one will be inclined to think that the inferences involving (1) must be explainable in terms of (5). This is why now it is quite common to claim, against Frege, that the logical form of (1) does not substantially differ from that of (2).[3]

However, it might be argued that this view is not immune to troubles. If one assumes that the logical form of a sentence is determined by its LF, one will be unable to provide a formal explanation of all the logical relations in which the sentence may be involved. For some of those relations hold in virtue of the content expressed by the sentence. This turns out clear if we consider context sensitive sentences, which express different contents in different contexts even though their LF remains the same. To illustrate, consider (1) and the following sentence:

7. Not all philosophers are rich

Imagine that you utter (1) with the intention to assert that all philosophers in your university are rich, while I utter (7) with the intention to assert that some philosophers in my university are not rich. There is an obvious sense in which we are not contradicting each other. But if the formal representation of (1) and (7) does not take into account the content they express, the apparent absence of contradiction is not formally explained. For the formula assigned to (7) must be the negation of the formula assigned to (1). More generally, let Γ be a set of sentences such that some of its members contain context sensitive expressions. In order to provide a formal explanation of the logical relations in Γ, the formal representation of Γ must display the semantic relations between the contents expressed by the sentences

route. Let $\forall x\alpha$ be $\forall xPx$ and consider a variable y distinct from x. Do $\forall x$ and $\forall y$ denote the same function? On the one hand, it seems that they should. If two functions assign the same values to the same arguments, as it is presumable in this case, then they are the same function. On the other, however, it seems that they should not. If $\forall x$ and $\forall y$ have the same meaning, then their meaning must be combinable in the same way with the meanings of other expressions. But $\forall yPx$ does not have the same meaning as $\forall xPx$. As a matter of fact $\forall yPx$ is not even a sentence, so it cannot be evaluated as true or false.

[3]The formal approaches to natural language derive to a good extent from Montague [7]. The view that a unique syntactic notion of logical form is able to provide both a compositional account of meaning and a formal explanation of logical properties emerges in several recent works such as Neale [8], Stanley [11] and Borg [2].

in Γ. However, this is not possible if logical form is individuated in terms of LF. For according to such a criterion of individuation, the logical form of each of the sentences in Γ does not depend on the content it expresses. Arguably, this is a serious limitation, which prevents any syntactic notion of logical form from being ideal for the purpose of formal explanation.[4]

The misgivings considered suggest that neither of the two views is entirely satisfactory: on the one hand, it is not obvious how a compositional account of the meaning of quantified sentences can be provided by their representation in a classical first order language; on the other, it is not obvious how an adequate formal explanation of the logical relations involving quantified sentences can be provided in terms of their syntactic structure. Therefore, unless such misgivings are dispelled, it is reasonable to presume that there is something wrong with the uniqueness assumption that underlies both views, namely, the assumption that one and the same notion of logical form can provide answers to both (Q1) and (Q2).

In what follows it will be taken for granted that different notions of logical form can be employed to address (Q1) and (Q2). More specifically, the hypothesis that will be held about (Q1), which is the focus of this paper, is that the notion of logical form that suits the purpose of formal explanation is *truth conditional*, that is, it is a notion according to which logical form is determined by truth conditions. Since no uniqueness assumption about (Q1) and (Q2) will be adopted, this is compatible with there being a different notion of logical form that is suitable for (Q2). So it is compatible with the hypothesis that a syntactic notion of logical form is to be adopted to answer (Q2).

The truth conditional notion of logical form stems from the idea that an adequate formalization of a sentence s must provide a representation of what is said by uttering s. For what is said by uttering s cannot be represented unless the truth conditions of s are exhibited. Obviously, this does not mean that what is said by uttering s is reducible to the truth conditions of s, at least if truth conditions are understood as sets of possible worlds, and sameness of truth conditions is rendered as sameness of truth value in every structure. It is reasonable to presume that only some of the formulas that preserve the truth conditions of s in that sense adequately formalize s. For example, it is usually taken for granted that Fa is better than $\sim\sim Fa$ or $Fa \wedge (Gb \vee \sim Gb)$ as a representation of 'Fido is a dog': even though $\sim\sim Fa$ and $Fa \wedge (Gb \vee \sim Gb)$ have the same truth value as Fa in every structure, they do not capture what is said by using 'Fido is a dog' in the relevant sense of 'what is said'. Nonetheless, preservation of truth conditions may plausibly be regarded as a necessary condition of adequate formalization.[5]

[4] Iacona [5] provides a more articulated defence of this claim.

[5] Sainsbury [10] suggests a criterion of adequate formalization that rests on the idea that formalization must preserve what is said, pp. 161–162.

7.2 Some Terminology

In order to provide a formal account of quantified sentences based on a truth conditional notion of logical form, a principled distinction must be drawn between the meaning of quantified sentences and their truth conditions. This section introduces some terminology that will be employed to phrase the distinction.

In the first place, the term 'quantifier expression' will be used to refer to expressions such as 'all' or 'some', which occur in noun phrases as determiners of nominal expressions. In accordance with this use, we will restrict attention to simple quantified sentences that contain expression of this category, such as (1) or the following:

8. Some philosophers are rich

In the second place, the term 'domain' will be used to refer to the totality of things over which a quantifier expression is taken to range. In ordinary talk, quantifier expressions often carry a tacit restriction to a set of contextually relevant objects. For example, on one occasion (1) may be used to assert that all philosophers in a university U are rich, while on another occasion it may be used to assert that all philosophers in another university U' are rich. So it is presumable that in the first case 'all' ranges over a set of people working or studying in U, while in the second it ranges over a set of people working or studying in U'. In order to take into account contextual restrictions of this kind it will be assumed that, whenever a quantifier expression is used, some domain is associated with its use, that is, the domain over which the quantifier expression is taken to range.[6]

In the third place, the term 'quantifier' will be used to refer to functions from domains to binary relations. In accordance with this use, the meaning of 'all' may be defined as a quantifier *all*, that is, as a function which, for any domain D, denotes a binary relation that satisfies the following condition:

Definition 7.1. $all_D(A,B)$ if and only if $A \subseteq B$.

Here A and B are sets whose members belong to D, and the left-hand side is read as 'the relation denoted by 'all' relative to D obtains between A and B'.

The meaning of 'some' may be defined in similar way as a quantifier *some*, that is, as a function which, for any D, denotes a binary relation that satisfies the following condition:

[6]This assumption leaves unsettled the question of how the restriction is determined in the context. More specifically, it is neutral with respect to the divide between semantic and pragmatic accounts of domain restriction. The accounts of the first kind represent domains by some sort of parameters in the noun phrase, either in the determiner or in the noun. Those of the second kind, instead, leave the determination of domains to pragmatic factors which determine the communicated content as distinct from what is literally said.

Definition 7.2. $some_D(A, B)$ if and only if $A \cap B \neq \emptyset$.[7]

The relativization to domains involved in Definitions 7.1 and 7.2 accounts for the fact that the extension of a quantifier expression may vary from occasion to occasion, even though its meaning does not change. If e is a quantifier expression that means Q, then Q_D is the extension of e relative to D. Thus if D is a set of people working or studying in U and D' is a set of people working or studying in U', 'all' denotes different relations relative to D and D'. So there is a sense in which 'all' means the same thing on both occasions, yet the relations denoted differ. The same goes for 'some'. More generally, a distinction may be drawn between *global* quantifiers and *local* quantifiers, that is, between quantifiers as functions from domains to binary relations and quantifiers as values of such functions. If Q is a global quantifier and D is a domain, then Q_D is the local quantifier assigned by Q to D.[8]

If the meaning of quantifier expressions is defined in the way outlined, and it is assumed that nominal expressions denote sets, the meaning of quantified sentences is easily obtained by composition. Let A and B be sets denoted by 'philosophers' and 'rich' relative to D. For example, if D is a set of people working or studying in U, A and B are subsets of that set. Given Definition 7.1, all_D fixes truth conditions for (1) relative to D, that is, (1) is true if and only if $A \subseteq B$. So the meaning of (1) may be described as a function from domains to truth conditions, which results from the combination of *all* with the meanings of 'philosophers' and 'rich'. The case of (8) is similar. Assuming that A and B are sets denoted by 'philosophers' and 'rich' relative to D, the meaning of (8) may be described as a function from domains to truth conditions which results from the combination of *some* with the meanings of 'philosophers' and 'rich'. More generally, the meaning of a quantified sentence s that contains a quantifier expression e that means Q is a function from domains to truth conditions that is obtained by combining Q with the meaning of the nominal expressions in s. The value of the function for each D is determined by Q_D, that is, by the local quantifier assigned by Q to D.

7.3 Formalization and Interpretation

Section 7.2 shows how a principled distinction can be drawn between the meaning of quantified sentences and their truth conditions. The meaning of a quantified sentence s results from the composition of the meanings of its constituent expressions, so it belongs to s independently of how s is understood on this or that occasion. The truth conditions of s, instead, are fixed by the domain associated with the quantifier

[7] Definitions 7.1 and 7.2 are as in Peters and Westerståhl [9], pp. 62–64.

[8] The distinction between global quantifiers and local quantifiers is drawn in Peters and Westerståhl [9], p. 48.

expression that occurs in s, so they depend just on how s is understood on this or that occasion.

Let an *interpretation* of a sentence be an assignment of semantic properties that determines definite truth conditions for the sentence in accordance with the meaning of its constituent expressions. On the formal account of quantified sentences that will be suggested, quantified sentences have logical form relative to interpretations. For interpretations fix domains for the quantifier expressions occurring in them.

The hypothesis that will be adopted is that quantified sentences can be formalized in L by means of formulas that represent their truth conditions relative to interpretations. To illustrate, consider (1). The simplest way to represent (1) in L is by means of (3). The representation provided by (3) includes no restriction on the domain. Note that the assumption that quantifier expressions are used in association with domains does not entail that, whenever one uses a quantifier expression, one has in mind a set of contextually relevant objects. It is consistent with that assumption to say that there are contexts in which nothing is excluded as irrelevant. So (3) represents (1) as used in such a context. In other words, (3) represents the truth conditions of (1) relative to an interpretation whose domain is the totality of everything.

In order to deal with a context in which some things are excluded as irrelevant, the intended restriction may be stated as part of the formula. Suppose that (1) is used to assert that all philosophers in U are rich. In this case, (1) may be represented as follows:

9. $\forall x(Rx \supset (Px \supset Qx))$

Here R stands for a condition that applies to a set of people working or studying in U. So if two utterances of (1) differ in the intended restriction on the domain, they may be represented by means of different predicate letters. Suppose that (1) is used in one context to assert that all philosophers in U are rich and in another context to assert that all philosophers in U' are rich. This difference may be represented in terms of the difference between (9) and the following formula:

10. $\forall x(Sx \supset (Px \supset Qx))$

Here S stands for a condition that applies to a set of people working or studying in U'. From (9) and (10) it turns out clear that (1) has different truth conditions relative to different interpretations. Note that if (1) and (7) are formalized in this way, the example considered Sect. 7.1 can easily be handled as a case where no contradictory pair of formulas is involved.

The case of (8) is similar. The simplest way to represent (8) in L is the following:

11. $\exists x(Px \wedge Qx)$

Again, this representation includes no restriction on the domain. In order to deal with a context in which some things are excluded as irrelevant, the intended restriction may be stated as part of the formula. From now on, however, considerations about restricting conditions will be avoided for the sake of simplicity.

7.4 The Issue of First Order Definability

A major implication of the thesis that quantified sentences can be formalized in L in virtue of their truth conditions concerns a fact that is usually regarded as decisive for the issue of the expressive power of classical first order logic. The fact is that some quantifier expressions are not *first order definable*, in the sense that they do not denote quantifiers that satisfy the following condition:

Definition 7.3. A quantifier Q is *first order definable* if and only if there is a formula α of L containing two unary predicate letters such that, for every set D and $A, B \subseteq D$, $Q_D(A, B)$ if and only if α is true in a structure with domain D where the predicate letters in α denote A and B.

As it is easy to verify, 'all' is first order definable, because (3) is a formula of L containing two unary predicate letters such that, for every set D and $A, B \subseteq D$, $all_D(A, B)$ if and only if (3) is true in a structure with domain D where its predicate letters denote A and B. The same goes for 'some', given that (8) can be represented as (11).

However, not all quantifier expressions are like 'all' and 'some'. Consider the following sentence, which contains the quantifier expression 'more than half of':

12. More than half of philosophers are rich

The quantifier *more than half of* may be defined as a function which, for any D, denotes a binary relation that satisfies the following condition:

Definition 7.4. $\textit{more than half of}_D(A, B)$ if and only if $| A \cap B | > 1/2 \, | A |$

Although this definition differs from Definitions 7.1 and 7.2 in that it involves a proportional relation that applies to the cardinality of A and B, *more than half of* is a function from domains to binary relations exactly like *all* and *some*. So (12) is semantically similar to (1) and (8), in that it is formed by expressions of the same semantic categories combined in the same way. However, there is no formula of L that translates (12) in the same sense in which (3) and (11) translate (1) and (8). This is to say that 'more than half of' is not first order definable.[9]

Many are inclined to think that this fact constitutes a serious limitation of the expressive power of first order logic. If it is assumed that formalization is a matter of translation, understood as meaning preservation, then it is natural to think that there is no way to formalize (12) in L. More generally, one may be tempted to think that a quantified sentence can be formalized in L only if the quantifier expressions it contains are first order definable.[10]

Without that assumption, however, there is no reason to think that the first order undefinability of more than half of' rules out the possibility that (12) is formalized

[9]Barwise and Cooper [1], pp. 213–214, provides a proof of the first order undefinability of 'more than half of'.

[10]As in Barwise and Cooper [1], p. 159.

in L. Certainly, it undermines the claim that there are sentences of L that have the same meaning as (12). But if logical form is determined by truth conditions, such a claim makes little sense anyway, even in the case of (1) and (8). For formalization is not a matter of translation, but a matter of representation of truth conditions.

Instead of asking whether a quantifier expression is first order definable, one may ask whether it is *first order expressible*, that is, whether it denotes a quantifier that satisfies the following condition:

Definition 7.5. A quantifier Q is *first order expressible* if and only if, for every set D and $A, B \subseteq D$, there is an adequate formula α of L containing two unary predicate letters such that $Q_D(A, B)$ if and only if α is true in a structure with domain D where the predicate letters denote A and B.

The sense in which α is required to be adequate is the same sense in which a formalization is expected to be adequate, as explained in Sect. 7.1: α must represent what is said, relative to D, by a sentence which contains a quantifier expression that denotes Q and two predicates for A and B. Clearly, adequacy so understood cannot be phrased in formal terms, as the notion of what is said is irreducibly vague. However, the condition that α is adequate is clear enough for the purposes at hand, or so it will be assumed.

To see how adequacy matters, it suffices to think that a trivial proof of the existence of α can easily be provided if no such condition is imposed on α. For it easy to find some α that has the required truth value in the structure for independent reasons. For example, if $Q_D(A, B)$ and α is a logical truth, then $Q_D(A, B)$ if and only if α is true in the structure. However, it is clear that in this case α is not adequate. The same goes for similar trivial proofs of the existence of α. What is not trivial, instead, is to prove the existence of an adequate α. As it will be shown, 'more than half of' is first order expressible, in that for every D and $A, B \subseteq D$, there is an adequate sentence α of L containing two predicate letters such that *more than half of*$_D(A, B)$ if and only if α is true in a structure with domain D where the predicate letters denote A and B.

The proof that will be provided rests on two assumptions. The first is that A and B are finite. This is an assumption that one can plausibly make when one restricts attention to natural language, for 'more than half of' is normally used to state relations between finite quantities, as indicated by the proportion $1/2$ that occurs in Definition 7.4. This is not to deny that 'more than half of' can be used in some intelligible way for infinite domains. Presumably, some technical or semi-technical meaning can be specified for that purpose. However, infinitary uses of 'more than half of' will not be considered in what follows. Independently of how such uses relate to the ordinary understanding of the expression, the reasoning simply will not apply to them.[11]

The second assumption is that, if what is said by s relative to D is that at least n As are Bs, then a formula of L that contains n occurrences of $\exists$ and two unary

[11] Barwise and Cooper [1], p. 163, consider infinitary uses of 'more than half of'.

predicates P and Q can provide an adequate representation of s. More precisely, let the symbol $\exists_{\geq n}$ be used to abbreviate formulas of L in the following way: $\exists_{\geq n}\bar{x}\alpha(\bar{x})$ means $\exists x_1 \ldots \exists x_n(\alpha(x_1) \wedge \ldots \wedge \alpha(x_n) \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j)$, where $\alpha(x_i)$ is a formula in which x_i occurs free, and in the second part of the conjunction every x_i is said to differ from every other. Then, if what is said by s relative to D is that at least n As are Bs, then the following formulas adequately represents s:

13. $\exists_{\geq n}\bar{x}P(\bar{x}) \wedge Q(\bar{x})$

For example, suppose that D includes some persons, and that three of them are philosophers. Then what is said by (12) relative to D is that at least two philosophers are rich, which is adequately represented by the formula $\exists x \exists y(Px \wedge Qx \wedge x \neq y)$.

Given these two assumptions, the first order expressibility of 'more than half of' can be proved in two steps.

Theorem 7.1. *If $A, B \subseteq D$, there is an n such that $\mid B \mid > 1/2 \mid A \mid$ if and only if $\mid B \mid \geq n$.*

Proof. Let F be a function defined as follows. If $m = 0$, then $F(m) = 1$. If $m > 0$ and m is even, then

$$F(m) = \frac{m+2}{2}$$

If $m > 0$ and m is odd, then

$$F(m) = \frac{m+1}{2}$$

Let $\mid A \mid = m$ and $n = F(m)$. n is such that $\mid B \mid > 1/2 \mid A \mid$ if and only if $\mid B \mid \geq n$. Suppose that $m = 0$. Then $1/2 \mid A \mid = 0$ and $F(m) = 1$, so $\mid B \mid > 0$ if and only if $\mid B \mid \geq 1$. Suppose that $m > 0$ and m is even. Then there is a k such that $m = 2k$, hence $\mid B \mid > 1/2 \mid A \mid$ if and only if $\mid B \mid > k$. Moreover,

$$F(m) = \frac{m+2}{2} = \frac{2k+2}{2} = \frac{2(k+1)}{2} = k+1$$

Therefore, $\mid B \mid > k$ if and only if $\mid B \mid \geq k+1$. Finally, suppose that $m > 0$ and m is odd. Then there is a k such that $m = 2k+1$, hence $\mid B \mid > 1/2 \mid A \mid$ if and only if $\mid B \mid > k + 1/2$. By hypothesis, $\mid B \mid$ is a natural number, so $\mid B \mid > k + 1/2$ if and only if $\mid B \mid > k$. Moreover,

$$F(m) = \frac{m+1}{2} = \frac{2k+1+1}{2} = \frac{2(k+1)}{2} = k+1$$

Therefore, $\mid B \mid > k$ if and only if $\mid B \mid \geq k+1$.

Theorem 7.2. *For every D and $A, B \subseteq D$, there is an adequate sentence α of L that contains two unary predicate letters such that* more than half of$_D(A, B)$ *if and only if α is true in a structure with domain D where the predicate letters denote A and B.*

Proof. Let $A, B \subseteq D$. From Theorem 7.1, replacing B with $A \cap B$, it turns out that there is an n such that $| A \cap B | > 1/2 | A |$ if and only if $| A \cap B | \geq n$. By Definition 7.4, there is an n such that *more than half of*$_D(A, B)$ if and only if $| A \cap B | \geq n$. The condition that $| A \cap B | \geq n$ is adequately expressed in L by (13). Moreover, (13) is true in a structure with domain D where P and Q denote A and B, and more than half the As are Bs.[12]

Theorem 7.1 expresses the obvious truth that, for every finite set, there is an n such that saying 'more than half of' amounts to saying 'at least n'. This guarantees that, although the global quantifier *more than half of* is characterized by a proportional relation, each local quantifier *more than half of*$_D$ fixes a non-proportional relation expressible in L. Theorem 7.2, accordingly, "squeezes" a proportional relation on a set of non-proportional relations. So we get that, for any domain, (12) has a logical form representable in L relative to that domain. This means that, for any interpretation, (12) has a logical form representable in L relative to that interpretation.

7.5 Generalization

The account of 'more than half of' suggested in Sect. 7.4 may easily be extended to other quantifier expressions whose meaning is definable in terms of proportional relations, such as 'most', 'few' and 'many'. Even though 'most', 'few' and 'many' exhibit a kind of indeterminacy that does not affect 'more than half of', in that they admit multiple admissible readings, this difference does not prevent them from being amenable to the same kind of treatment that applies to 'more than half of'.

To illustrate, let us focus on 'most'. A basic fact about its meaning seems to be that the condition stated in Definition 7.4 must be satisfied for the intended relation to obtain. Consider the following sentence:

14. Most philosophers are rich

If one utters (14), one says at least that more than half of philosophers are rich. However, this is a necessary but not a sufficient condition. Although 'most' may be used as synonymous of 'more than half of', its meaning seems to allow for variation in the proportion between the size of $A \cap B$ and the size of A. In order to account for this variation, a definition of *most* may be given along the following lines:

[12] The number triangle method outlined by Peters and Westerståhl in [9], pp. 160–161, provides a clear visual representation of the fact that *more than half* determines an n on every finite domain.

Definition 7.6. $most_D(A, B)$ if and only if $\mid A \cap B \mid > n/m \mid A \mid$

Here $0 < n < m$ and $n/m \geq 1/2$. For example, 1/2 and 2/3 are equally admissible values for n/m. In other words, *most* is defined as a class of quantifiers rather than as a single quantifier. Consequently, the meaning of (14) may be described as a class of functions from domains to truth conditions that is obtained by combining *most* with the meanings of ‘philosophers’ and ‘rich’. This means that (14) differs from (12), in that the determination of its truth conditions involves a parameter other than the domain. Let A and B be the sets denoted by ‘philosophers’ and ‘rich’ relative to D. Whether $most_D$ obtains between A and B depends on the values assigned to n and m. For example, if $n = 2$ and $m = 3$, then it obtains just in case $\mid A \cap B \mid > 2/3 \mid A \mid$. In order to determine definite truth conditions for (14), we need both a domain and a value of the additional parameter.[13]

If *most* is defined in the way suggested, the distinction between first order definability and first order expressibility drawn in Sect. 7.4 can be applied to (14). Although it is a fact that ‘most’ is not first order definable, on the assumption that logical form is determined by truth conditions (15) can be formalized in L independently of this fact. For what matters is that ‘most’ is first order expressible.[14]

To show that (14) can be formalized in L, it suffices to prove a squeezing result similar to Theorem 7.2. This can be done by means of a generalization of Theorem 7.1: if $A, B \subseteq D$ and $0 < n < m$, there is a k such that $\mid B \mid > n/m \mid A \mid$ if and only if $\mid B \mid \geq k$. From such generalization it follows that, for every D and $A, B \subseteq D$, there is an adequate sentence α of L that contains two unary predicate letters such that $most_D(A, B)$ if and only if α is true in a structure with domain D where the predicate letters denote A and B.

As in the case of ‘most’, the meanings of ‘few’ and ‘many’ may be defined as classes of quantifiers *few* and *many*. So it may be assumed that the meaning of the following sentences is obtained by combining *few* and *many* with the meanings of ‘philosophers’ and ‘rich’:

15. Few philosophers are rich
16. Many philosophers are rich

The meaning of (15) and (16) may thus be described as a class of functions from domains to truth conditions. This suggests that, as in the case of *most*, a squeezing argument can be provided to the effect that *few* and *many* are first order expressible.[15]

[13]Definition 7.6 is in line with the suggestion in Barwise and Cooper [1], p. 163, and the account in Westerståhl [12, pp. 405–06]. In the latter work, two readings of ‘most’ are considered. But if Definition 7.6 is adopted there seems to be no reason to do that.

[14]Peters and Westerståhl, in [9], pp. 466–468, outline a proof method that can be employed to show that ‘most’ and other proportional quantifiers are not first order definable.

[15]The case of ‘few’ and ‘many’ is definitely more controversial. For example, Keenan and Stavi [6] excludes that ‘few’ and ‘many’ can be treated in this way.

In substance, (14)–(16) can be treated in the same way as (12), with the only difference that in the case of (14)–(16) some parameter other than the domain must be taken into account as relevant to the determination of truth conditions. Therefore, on the assumption that an interpretation of (14)–(16) includes both a domain and a value for such a parameter, it turns out that, for every interpretation of (14)–(16), there is a formula of L that represents the truth conditions of (14)–(16) relative to that interpretation.

7.6 Logicality

The point that emerges from Sects. 7.4 and 7.5 is that it must not be assumed that first order definability is the property to be considered in order to settle the question whether quantified sentences can adequately be formalized in a classical first order language. On the formal account of quantified sentences suggested here, the property to be considered is first order expressibility. This does not mean, however, that first order definability is not a significant property. As it will be suggested, there is a straightforward relation between first order definability and *logicality*.

The quantifier expressions traditionally studied by logicians, such as ‘all’ or ‘some’, have always been regarded as paradigmatic examples of logicality. However, there are many more quantifier expressions than those traditionally studied by logicians. So it is natural to ask whether all quantifier expressions must be classified as logical. According to Barwise and Cooper they must not, in that there is no reason to think that the meaning of every quantifier expression is to be “built into the logic”. A distinction must be drawn between logical and non-logical quantifier expressions: ‘all’ and ‘some’ belong to the first category, while ‘more than half’, ‘most’, ‘many’ and ‘few’ belong to the second. The method of formalization adopted here provides one way to substantiate this distinction.[16]

We saw that, for every interpretation of a quantified sentence *s*, there is a formula of L that represents the truth conditions of *s* on that interpretation. Therefore, different formulas of L may represent *s* on different interpretations. But there are basically two ways in which the formal representation of *s* can vary as a function of its interpretation. Consider (1) and (12). In the case of (1), the variation concerns at most the non-logical vocabulary of L, as in (9) and (10). In the case of (12), instead, it may also concern the logical vocabulary of L. For example, the following formulas of L represent the logical form of (12) on different interpretations:

17. $\exists_{\geq 3}\bar{x}P(\bar{x}) \wedge Q(\bar{x})$
18. $\exists_{\geq 4}\bar{x}P(\bar{x}) \wedge Q(\bar{x})$

[16] Barwise and Cooper [1], p. 162.

One thing is to say that more than half of five things have a certain property, quite another thing is to say that more than half of six things have that property.

The contrast between the two cases considered may be described in terms of two kinds of variation in the formal representation of a sentence *s*. A *weak variation* in the formal representation of *s* depends on some difference in the non-logical vocabulary of the formulas assigned to *s*. Instead, a *strong variation* in the formal representation of *s* depends on some difference in the logical vocabulary of the formulas assigned to *s*. So, the first case may be described as one in which a difference between two interpretations entails weak variation in the formal representation of (1), as in (10) and (11), while the second may be described as one in which a difference between two interpretations entails strong variation in the formal representation of (13), as in (18) and (19).

There is a plausible sense in which weak variation, unlike strong variation, does not entail difference in logical form. This is to say that *sameness of logical form* may be understood in terms of weak variation: *s* has the same logical form on two interpretations if and only if the difference between them entails at most weak variation in the formal representation of *s*. Logicality may be defined in terms of sameness of logical form so understood:

Definition 7.7. A quantifier expression is *logical* if and only if every quantified sentence in which it occurs has the same logical form on all interpretations.[17]

From Definition 7.7 it turns out that 'all' is logical. For (1) has the same logical form on all interpretations, whether or not its formalization includes a restricting condition. The same goes for (8). By contrast, 'more than half of', 'most', 'many' and 'few' are non-logical, for (12) and (14)–(16) have different logical forms on different interpretations.

Note that the sense of 'logical' provided by Definition 7.7 is essentially relative, in that it depends on the choice of logical constants that underlies the language in which logical forms are expressed. On the assumption that logical forms are expressed in L, 'logical' is to be read as relative to L. This, however, should not be regarded as a flaw. Definition 7.7 is neutral with respect to the notoriously controversial question of whether an absolute criterion of logical constancy can be specified in non-circular way. If the answer to that question is affirmative, then it is presumable that some independent justification of the choice of logical constants that underlies L can be provided. If it is negative, instead, then the choice of logical constants that underlies L is itself in need of justification, so an account of logicality based on L is definitely circular. Even though it is arguable that only in the first case we can get an interesting distinction between logical and non-logical quantifier expressions, in any case the relativity involved in Definition 7.7 causes no trouble by itself.

[17]Note that, given the restriction mentioned in Sect. 7.2, 'quantified sentence' refers to simple quantified sentences such as (1) or (12). This rules out obvious counterexamples such as 'Most but not all philosophers are rich'.

There is a straightforward relation between logicality so defined and first order definability:

Theorem 7.3. *Every logical quantifier expression is first order definable.*

Proof. Let us assume that e is a logical quantifier expression that denotes a quantifier Q, and that s is a quantified sentence in which e occurs. Let α be a formula of L which contains two predicate letters and represents the truth conditions of s on some interpretation with domain D. Then it must be the case that, for $A, B \subseteq D$, $Q_D(A, B)$ if and only if α is true in a structure with domain D where the predicate letters in α denote A and B. Now take any domain D'. For some interpretation with domain D', there is a formula α' of L such that α' represents the truth conditions of s, so that, for $A', B' \subseteq D'$, $Q_{D'}(A', B')$ if and only if α' is true in a structure with domain D' where the predicate letters in α' denote A' and B'. But since e is logical, s has the same logical form on all interpretations. This means that α and α' differ at most in the predicate letters. Therefore, α' is true in a structure with domain D' where the predicate letters in α' denote A' and B' if and only if α is true in a structure with D' where the predicate letters in α denote A' and B'. This is to say that α satisfies the condition required by Definition 7.3, so that e is first order definable.

Theorem 7.3 characterizes logical quantifier expressions as first order definable quantifier expressions. This characterization entails that every quantifier expression that is not first order definable is not logical. So, the point that has been made in Sects. 7.4 and 7.5 may be refined as follows. Quantifier expressions such as 'more than half of', 'most', 'few' and 'many' are not first order definable. But this does not entail that the quantified sentences in which they occur cannot be formalized in a classical first order language. What it entails is at most that they are not logical.[18]

7.7 Conclusion

From the analysis of quantified sentences suggested in the previous sections it turns out that there is something right and something wrong in each of the two views considered in Sect. 7.1. On the one hand, there is a sense in which it is right to say that (1) and (2) are structurally different, namely, that in which (1) and (2) are adequately represented as (3) and (4) in order to formally explain the inferences involving them. On the other, there is a sense in which it is right to say that (1) and (2) are structurally similar, namely, that in which (1) and (2) are adequately represented as (5) and (6) in order to provide a compositional account of their meaning. What is wrong is to think that there must be a unique sense in which

[18]There is an interesting convergence between the account of logical quantifier expressions suggested here and the independently motivated account outlined in Feferman [3], see p. 140. As it is noticed in that work, pp. 144–145, it is not as obvious as it might seem that the converse of Theorem 7.3 is guaranteed to hold.

either (1) and (2) are structurally different or they are structurally similar. On the understanding of logical form that is suitable to address (Q1) they are structurally different, while on the understanding of logical form that is suitable to address (Q2) they are structurally similar. This is just another way of saying that there is no unique answer to the question of what is the logical form of quantified sentences.

References

1. Barwise, J., and R. Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4: 159–219.
2. Borg, E. 2007. *Minimal semantics*. Oxford: Oxford University Press.
3. Feferman, S. 2015. Which quantifiers are logical? A combined semantical and inferential criterion. In *Quantifiers, quantifiers, and quantifiers*, ed. A. Torza. Springer.
4. Frege, G. 1879/1967. Concept script, a formal language of pure thought modelled upon that of arithmetic. In *From Frege to Godel: A sourcebook in mathematical logic*, ed. J. van Heijenoort, chapter Begriffsschrift, 5–82. Cambridge: Harvard University Press.
5. Iacona, A. 2013. Logical form and truth conditions. *Theoria* 28: 439–457.
6. Keenan, E.L., and J. Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9: 253–326.
7. Montague, R. 1974. *Formal philosophy*. New Haven: Yale University Press.
8. Neale, S. 1993. Logical form and LF. In *Noam Chomsky: Critical assessments*, ed. C. Otero, 788–838. London: Routledge.
9. Peters, S., and D. Westerståhl. 2006. *Quantifier in language and logic*. Oxford: Oxford University Press.
10. Sainsbury, M. 1991. *Logical forms*. Oxford: Blackwell.
11. Stanley, J. 2000. Context and logical form. *Linguistics and Philosophy*. Oxford, 23: 391–434.
12. Westerståhl, D. 1985. Logical constants in quantifier languages. *Linguistics and Philosophy* 8: 387–413.

Chapter 8
Quantification with Intentional and with Intensional Verbs

Friederike Moltmann

Abstract Whether natural language permits quantification over 'nonexistent', intentional objects is subject of a major controversy, as is the nature of such entities themselves. This paper argues that certain constructions in natural language involving 'intentional verbs' such as '*think of*', '*describe*', and '*imagine*' cannot be analysed compositionally without positing intentional objects, as entities strictly dependent on intentional acts. The paper also argues that intentional verbs involve a distinctive semantics, which is fundamentally different from that of intensional transitive verbs, a difference reflected in a range of quantificational phenomena.

The questions whether natural language permits quantification over intentional objects as the 'nonexistent' objects of thought is the topic of major philosophical controversy, as is the status of intentional objects as such. Many philosophers deny the possibility of there being 'nonexistent' objects of thought. Others following Meinong [12], take 'nonexistent' objects of thought to be entities individuated only by a particular set of properties, and as having a weaker form of being than existence. Yes others, in the tradition of Brentano [1], admit the possibility of intentional, nonexistent objects, but take them to be dependent on an intentional act or state. This paper will argue that natural language does reflect a particular notion of intentional object and in particular that certain types of natural language constructions (generally disregarded in philosophical literature) cannot be analysed without positing intentional objects. At the same time, those intentional objects do not come for free; rather they are strictly dependent on intentional acts that generally need to be present, in one way or another, in the semantic structure of the sentence.

The constructions in question display a particular dependence of intentional objects on the event argument of an intentional verb in the same sentence, a verb like

F. Moltmann (✉)
Institut d'Histoire et de Philosophie de Sciences et Techniques (IHPST), Université Paris 1, 13 rue du Four, 75006 Paris, France
e-mail: Friederike.Moltmann@univ-paris1.fr

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_8

think about, refer to, describe, or *imagine*. Intentional verbs, I will argue, involve a semantics different from that of extensional verbs and their semantics needs to be sharply distinguished from that of intensional verbs, verbs like *need*, *look for* and *owe*. Intentional and intensional verbs differ in a range of semantic properties, in particular in regard to quantificational complements.

The intentional objects that natural language involves are part of the domain of quantifiers and act as semantic values of referential terms, but as entities they are mere projections of what I call *quasi-referential acts*, namely either unsuccessful intentional acts or pretend acts of reference. Intentional objects depend for existence on quasi-referential acts and can bear (ordinary) properties only relative to acts of attribution. Moreover, what matters for the identity of intentional objects are relations of coordination among quasi-referential acts: two intentional objects are identical if the acts on which they depend are coordinated, regardless of what properties the entities are attributed.

The paper will first discuss the relevant natural language constructions and outline the required notion of an intentional object, proposing a particular semantic analysis of intentional verb constructions. Second, it will then contrast the semantics of intentional verbs with that of intensional verbs and address a range of further issues regarding the use of intentional objects in the analysis of natural language.

8.1 Quantification over Intentional Objects in the Context of Intentional Verbs

8.1.1 Intentional Objects in the Semantics of Natural Language

There is hardly a general agreement among philosophers that intentional or fictional object need to be posited, for the semantics of natural language or other purposes. One common strategy for avoiding fictional or intentional objects consists of making just use of quasi-referential acts and the relation coordination among them.[1] There are, however, constructions in natural language for which intentional objects not only appear to provide a straightforwardly analysis, but that could hardly be analysed compositionally without them. These are not the usual constructions discussed in the philosophical literature, though. The philosophical literature focuses on simple negative existentials as in (1) (consisting of a proper name or definite NP and an existence predicate such as *exist*) and sentences with transitive intentional verbs of the sort in (2)[2]:

---

[1] For such a view see Everett [4], Walton [36], and Taylor [29].

[2] Throughout this paper I assume that *exist* is a predicate. See Miller [13, 14] and Salmon [27, 28] for a philosophical defense of that view as well as Moltmann [20] for further linguistic considerations.

1. (a) The golden mountain does not exist.
   (b) Vulcan does not exist.
2. (a) John thought of the golden mountain.
   (b) John imagines a beautiful castle in the sky.

While intentional objects may provide a straightforward account of (1) and (2) preserving the uniformity of the semantics of definite and indefinite NPs, such sentences have hardly convinced philosophers in general of the necessity of positing intentional objects. A prominent approach to negative existentials as in (1), defended by Salmon [27, 28] as well as Sainsbury [26], is to take the subject of a negative existential to have an empty denotation and negation to be external. On that view, (1a) is to be understood as 'it is not true that the golden mountain exists', denying the truth of the sentence *the golden mountain exists*, rather than asserting its falsehood. Also the sentences in (2) do not seem to require intentional objects. The complements of verbs like *think of* or *imagine* could be taken to be that of intensional verbs, not requiring an actual referent, or as acting 'adverbially', as on adverbial approaches to intentionality.[3]

However, there are natural language constructions whose compositional analysis can hardly do without intentional objects. The *there*-sentence below is an example:

3. (a) There is a woman John is thinking about that does not exist.

In (3a), the object of John's thought belongs to the range of objects that the *there*-construction quantifies over, but it is an object that the existence predicate *exist* is not true of.

Here are further examples making the point ((3d) being a negative existential with the existence predicate *happen*):

(b) There are several buildings described in the catalogue that do not exist.
(c) There are two buildings mentioned on the map that do not exist.
(d) There is an accident John was thinking about that did not happen.

What is crucial in those examples is the occurrence of the *intentional verb* in the relative clause, that is, a transitive verb describing a mental act or speech act directed toward something possibly nonexistent. Without it, the sentences can hardly be considered true:

4. (a) There is a woman that does not exist.
   (b) There are several buildings that do not exist.
   (c) There is an accident that did not happen.

Sentences of this sort pose problems for certain Meinongian views on which nonexistent objects are mind-independent objects constituted by a (noninstantiated) set of properties.[4]

[3] See Tye [31].

[4] Such sentences are not a problem, though, for the Meinongian theory of Priest [24], who takes nonexistent objects to have 'existence-entailing' properties (such as the property of being a woman)

Quantification over intentional objects can also be achieved with two other sorts of constructions. One of them involves quantificational NPs, modified by relative clauses containing an intentional verb, as in (5a,b):

5. (a) Some women John mentioned do not exist.
   (b) Many buildings that John had planned never came into existence.

The other one involves quantificational NPs that are complements of intentional verbs:

6. (a) John mentioned some woman that does not exist.
   (b) Mary had described a building that never had come into existence.
   (c) Mary made reference to a poet that does not exist.
   (d) The book is about a detective that does not exist.

Intentional verbs allow for the introduction of intentional objects both as main verbs and in relative clauses. Intentional verbs are not the only linguistic means, though, of making intentional objects available for quantification. In addition adjectival modifiers such as *imaginary* enable quantification over intentional objects:

7. There are imaginary women that do not exist.

Of course, also the relational-noun construction *object of thought* itself can be used for that purpose[5]:

8. There are objects of thought/objects of imagination/objects of fantasy that do not exist.

Furthermore nouns like *topic* and *subject matter* enable quantification over intentional objects[6]:

---

only in other possible worlds, the worlds that realize the content of the fiction or the relevant intentional acts or states. It seems that this account makes predictions about modal statements, though, that are unsupported by linguistic intuitions. Sentences such as (ia) below do not seem any better than (ib):

(i) (a) There is something that could be a tree that does not exist.
    (b) There is a tree that does not exist.

[5]The noun *object* in the construction 'object of thought' is in fact a relational noun since it cannot be replaced by a noun like *entity* or *thing*. *Object of thought* describes whatever entities may stand in the object of-relation to a thought or other intentional state or act, be it a real object of some type or a 'nonexistent' object. See also Crane [3].

[6]Also the non-sortal noun *thing* allows for quantification over intentional objects in *there*-sentences:

(i) There are things that John imagined/thought about/made reference to that do not exist.

In this function, it need not match gender features of the intentional object:

(ii) There is something John was thinking about, a son who would one day take over his company.

9. (a) There are many topics John talked about, the woman Bill had mentioned, the pet Joe dreamt about etc.
   (b) There is a subject matter we did not discuss, namely the house John plans to build.

While the construction *object of thought* involves generic reference to intentional acts, nouns like *topic* and *subject matter* don't involve reference to intentional acts, at least not overtly.

8.1.2 The Notion of an Intentional Object

Constructions with intentional verbs display a particular notion of an intentional object as an object strictly dependent on an intentional act. Here a terminological distinction is needed between 'object of thought' and 'intentional object'. The object of a thought is what the thought is directed toward, which may be a real object or an object that does not exist, that is, a merely intentional object. In case a thought is not directed toward a real object but a merely intentional one, I will call the latter simply an 'intentional object', thus distinguishing—in a nonstandard way—the term 'object of thought' (which expresses a function that objects, real or nonexistent, may have) from the term 'intentional object' (as a term for a certain sort of nonexistent entity). Thus, for any object-directed attitude that is not directed towards a real object, there will be a corresponding intentional object.

Even though real and intentional objects may both play the role of objects of thought, they are not on a par otherwise. A merely intentional object is one that does not 'exist', and it may be an indeterminate or an impossible object.[7] Positing intentional objects thus does not mean taking unsuccessful acts of reference to in fact be successful, referring to intentional objects. Rather intentional objects are 'pseudo-objects' entirely constituted by unsuccessful or pretend acts of reference itself (and acts they are coordinated with). The non-existence of intentional objects thus is 'essentially and constitutively failed intentionality', as McGinn puts it (McGinn [11, p. 43]). Intentional objects are not part of the ontology; they are mere projections of intentional acts, which is why they have the status of nonexistents. Intentional objects thus are not peculiar types of objects that are by nature nonexistent.[8]

Intentional objects are dependent on intentional acts in a particularly strict way. An intentional object *o* dependent on an intentional act *a* does not have an (ordinary)

---

[7] One might posit the same type of intentional object for the two cases of attitudes and, in the case of an attitude being directed toward a real object, allow an intentional act to be related to two sorts of objects simultaneously: an intentional object and a real one. However, such a move is notoriously problematic: an intentional act just cannot relate to two such objects at once: it has a single object. This is a common objection raised against Brentano, see Voltolini [34] for discussion.

[8] For a critique of intentional objects in that sense see van Inwagen [33].

property P as such, but only relative to an act a' coordinated with a such that a involves the attribution of P to o. Whether two intentional objects are the same does not depend on whether they are attributed the same properties in intentional acts but whether the acts they depend on are coordinated. I will return to the relevant notion of coordination among intentional acts later.

8.1.3 Intentional Objects and Fictional Objects

Intentional objects as nonexistent objects need to be distinguished from fictional objects as objects that come into being by an act of creation. Intentional objects are the objects of certain object-related attitudes and linguistic acts that fail to be successful or were not meant to be successful. A fictional object is an entity that is created by producing a piece of fiction, and as a creation it exists. A fictional object is an object that an author creates as something going along with the story he is writing. A fictional object as a created object in that sense is an existent object, not a nonexistent one. It is an abstract artifact, to use Thomasson's [30] term.[9]

Intentional objects and fictional otherwise share important features. Both may be underspecified for a range of properties and be attributed contradictory properties in different acts (and even in a single act).[10] Both depend on intentional acts, in particular coordinated acts. But whereas fictional objects as abstract artifacts are true objects, intentional objects are quasi-objects: they are mere projections from unsuccessful or pretend referential acts. They in general can bear a property only relative to an act involving the attribution of that property, namely the object an associated referential act is meant to refer to.[11]

[9] The particular conditions that may distinguish a fictional object from an intentional object are further discussed in Thomasson [30] and Voltolini [35].

[10] The underspecification of intentional objects should not be confused with the nonspecificity of the complement of intensional transitive verbs, a point that will be discussed later.

[11] It is customary in the philosophical literature on fiction to distinguish between 'internal predication' and 'external predication'. *Sherlock Holmes lives on Baker Street* is true because the property of living on Baker Street is predicated of Holmes internally, whereas *Sherlock Holmes is a fictional character* is true because the property of being a fictional character is predicated of Sherlock Holmes externally. This suggests that intentional objects are predicated properties only internally, whereas fictional objects can be predicated properties both externally and internally. However, I think this distinction is not helpful. Intentional objects simply do not have properties as such, internally or externally predicated, rather they have properties only relative to an intentional act (including the intentional acts that make up a context of fiction).

A related distinction is Meinong's distinction between nuclear predicates and extranuclear predicates (see also Parsons [23]). However, the distinction between two ways of predicating is a better one, since one and the same predicate may be both internally and externally predicated. Another related distinction is the one van Inwagen [32] draws between having a property and holding a property. Having a property corresponds to a property externally predicated, whereas holding a property corresponds to a property internally predicated.

Intentional objects, as projections of quasi-referential acts, can bear certain properties, such as the property of being a topic of conversation, the property of being the object of an object-directed attitude, and of course the property of not existing. But intentional objects are not in space or time, whereas fictional objects as abstract artifacts have a temporal lifespan (Thomasson [30]).[12]

Given the distinction between fictional and intentional objects, a fictional name can be used in three different ways: for pretend reference (the referential use within a pretend context), as standing for an intentional object; and as standing for a fictional object. Only the fictional object can be the referent of a more complex term, of the sort the *fictional character Hamlet*, with its specific sortal *fictional character*.

The basis for intuitions about fictional objects is different from that for intuitions about intentional objects. Intuitions about fictional objects are based on our talk about fiction. By contrast, for at least the present purposes, intuitions about intentional objects will involve natural language constructions, in particular those with intentional verbs.

8.1.4 Intentional Objects and the Relation of Coordination Among Referential Acts

Intentional objects can be shared by different agents and by different acts of the same agent. Whether different acts or states share an intentional object depends not so much on whether the acts attribute the same properties to the object, but whether they are coordinated, that is, when the one act is meant to refer to or pretends to refer to the same object as the other act. Coordination among referential acts is an asymmetric relation among acts and to be understood roughly as follows: an act a is coordinated with an act a' iff a is meant to refer (or to pretend to refer) to the same object as a'.[13] Intentional acts of the same agent and of different agents may be coordinated.

The possibility of the same intentional objects to be shared by different, coordinated intentional acts is well-reflected in natural language. Different intentional objects can be the arguments of several intentional predicates, as long as the described intentional acts are coordinated, as in the examples below:

10. (a) John mentioned the woman the book is about.
    (b) John is thinking about the woman Mary told him about.

[12]What is described in a piece of fiction can also be viewed as an intentional object rather than a fictional object, namely as the intentional object that corresponds to the coordinated intentional acts that make up the writing of the fiction. It is the object the fiction is about, but it is not the object the author intended to create. The same pretend acts of reference thus give rise to two distinct objects: nonexistent intentional objects and fictional objects. The two kinds of objects may share the same internally predicated properties, but they differ in ontological status: one of them is a quasi-object, the other one is an abstract artifact.

[13]The relation of coordination as a relation among intentional acts thus differs from that of Fine [7], which is viewed primarily a relation among occurrences of expressions.

(c) John described the palace he had imagined.
(d) Jane told me about the woman John had described.

Intentional objects cannot be shared when there is no coordination among the relevant intentional acts, even if the respective acts involve the attribution of the same properties. Moreover, intentional objects can be the same even if they depend on acts attributing incompatible properties to them. Thus, (10a) would be true even if what John says about the woman contradicts what the book says about her.

Coordination among intentional acts is also reflected in the applicability of *the same* N. For two intentional acts to share 'the same N', it does not suffice that they involve intentional objects to which the same properties have been attributed (which is not even a necessary condition); rather the intentional acts (with respect to the relevant property attributions) have to have been coordinated—unless of course the objects of the acts are real objects. Consider the inference from (11a) and (11b) to (11c) and from (12a) and (12b) to (12c):

11. (a) John imagined a blue circle.
    (b) Bill imagined a blue circle.
    (c) John and Bill imagined the same circle.
12. (a) John thought of a tall woman with red hair.
    (b) Bill thought of a tall woman with red hair.
    (c) John and Bill thought of the same woman.

The inference in (11) is invalid, unless 'the same circle' is understood as 'the same type of circle', and similarly for (12) (assuming that the women in question do not exist), unless of course the intentional acts are coordinated.

The noun phrase *the same N*, for an ordinary noun N, needs to be sharply distinguished from the noun phrase *the same thing*, which gives rise to very different intuitions. The inference from (11a) and (11b) to (13a) is in fact valid, as is the inference from (12a) and (12b) to (13b):

13. (a) John and Bill imagined the same thing.
    (b) John and Bill thought of the same thing.

However we will see that *the same thing* involves an entirely different semantics than *the same N*, for an ordinary noun N: with *the same thing* in (13a,b), *imagine* and *think of* are used as intensional verbs, not intentional verbs, and *the same thing* in such sentences does not serve to express the sharing of an intentional object, but rather the sharing of a different, more abstract kind of entity.

The use of anaphora also reflects the importance of coordination for the identity of intentional objects:

14. (a) John described a castle. Bill described it too.
    (b) John dreamt about an extraordinary country. Bill dreamt about it too

As long as no real objects are involved, (14a) and (14b) imply that John's and Bill's acts of describing and dreaming are coordinated (which includes being related to a common source, for example a representation John and Bill both saw).

Thus, the identity of intentional objects does not so much depend on what properties they are attributed, but on the acts they depend on and on relations of coordination those acts enter with other acts. If John describes the woman he read about, John's description shares its intentional object with the one the book is about, as well as the act of writing the book. But John's description need not ascribe the very same property to that intentional object. John may remember the woman described incorrectly and the author may not be able to present the intended character well. Similarly, several acts of imagination may be about the same intentional object, involving the attribution of different properties. Of course, also different kinds of intentional acts may be coordinated. An act of describing may be coordinated with an act of imagination. We can then state the following conditions on intentional objects:

Conditions on intentional objects

15. (a) For any quasi-referential act e, there is an intentional object $f(e)$ of e.
(b) For quasi-referential acts e and e', the intentional object of e = the intentional object of e' iff e and e' are coordinated.
(c) For a quasi-referential act e, for an existence predicate E suited for the sortal that some act coordinated with e attributes to $f(e)$, E is false of $f(e)$.

The condition that the existence predicate be suited for the sortal property attributed to an intentional object accounts for the observation that, for example, *exist* can apply only to material and abstract objects and not events, whether actual or intentional, whereas an existence predicate like *happen* can apply to events only, whether actual or intentional (Moltmann [20]).

8.2 The Semantics of Intentional Verbs

8.2.1 The Interpretation of the Complement of Intentional Verbs

We have seen that the object position of transitive intentional verbs may involve intentional objects that are not generally available for quantification otherwise. The availability of intentional objects as entities in the domain of a quantifier obviously is tied to the intentional act described by the intentional verb. Given Davidsonian event semantics, this intentional act will be the Davidsonian event argument of the intentional verb.

The availability of intentional objects should not be made dependent directly, though, on the presence of an intentional event in the semantic structure of the sentence. This is because intentional objects can also be the semantic values of NPs like *subject matter* or *topic*. What is special about transitive intentional verbs is in fact not so much the described event making available an intentional object, but rather the ability of the complement to attribute a property constitutive of an intentional object, which is not the ordinary way of attributing a property to an object. This attribution of a property need not be part of the event described by the intentional verb though. If John thought of a detective, namely Sherlock Holmes, he need not know that Sherlock Holmes according to the fiction is a detective. But some coordinated act, at least that of the speaker uttering the sentence, attributes the property of being a detective to the object of John's thought.

The complement of an intentional verb is thus not evaluated in isolation, but relative to the event described by the verb. For present purposes, I will capture this dependence by simply taking nouns and adjectives to be polysemous, displaying an additional lexical variant involving an argument position for acts of attribution. Thus, for a 'non-relational' noun N, 'N(x, e)' is to be understood as 'some act coordinated with e attributes the property expressed by N to x'.

Note that the very same interpretation of the intentional verb—complement relation applies if the object of the intentional act or state turns out to be an actual object. This is important because the semantic interpretation of sentences with intentional verbs should stay neutral as to whether the described intentional act or state is successful or not. The semantics should not distinguish the case in which the complement of the intentional verb stands for an actual object from the one in which it stands for a merely intentional one. The complement can in both cases, for example, act as antecedent of anaphora in subsequent sentences.

The interpretation of the construction 'intentional verb-complement' will be based on coindexing of the nominal with the verb as below:

16. (a) $V_i \, DN'_i$

This syntactic relation then is interpreted by making use of the event argument of the verb for the interpretation of the nominal. The interpretation of a sentence like (16b) will thus be as in (16c):

(b) John mentioned$_i$ [a woman$_i$]
(c) ∃e ∃x(mention(e, John, x) & woman(x, e))

(16c) is to be understood as 'There is an event of mentioning an object on the part of John coordinated with an event of attributing the property of being a woman to that object'.

8.2.2 Relative Clauses with Intentional Verbs, with Past Tense and with Modals

The semantic analysis of intentional verbs cannot as yet apply to constructions with relative clauses containing an intentional verb as below:

17. (a) The woman John has described

The problem is that the standard compositional semantics of relative clause constructions cannot apply to this construction: the head noun would have to take as one of its arguments the event described by the intentional verb inside the relative clause, which is impossible. However, there is a syntactic view according to which the head of the relative clause originates from the lower position inside the relative clause. More specifically, it has been argued that the head noun of a relative-clause construction may originate from inside the relative clause.[14] If moreover movement of an expression is in fact copying, then an unpronounced copy of the expression moved will be left behind which will then be the one taken into account by semantic interpretation. This means that the construction can be interpreted as if the head noun was in the lower position, either by having left a copy in that position (on the copy theory of movement) or by being reconstructed into the lower position. The copy left behind should have the status of a restricted variable, bound by a silent operator that stands for the relative pronoun. The syntactic issues need not concern us in detail. Rather it suffices to take advantage of the general syntactic view that permits *the woman John described* be interpreted as below:

(b) The e [that John described e woman]
(c) $\iota x[\exists$ e(mention(e, John, x) & woman(x, e))]

Quantification over nonexistent objects can also be made available with modals and past tense which extend the range of a quantifier to past and possible objects. Modals and past tense pattern exactly the same as intentional verbs in relative-clause constructions:

18. (a) There are buildings that John could have built that do not exist.
    (b) There are many buildings built in the eighteenth century that do not exist anymore.
    (c) There are buildings that do not exist.

The relative clauses in (18a) and (18b) permit *there*-sentences to range over possible and past objects of which the existence predicate is not true, which is not the case for (18c), which can hardly be considered true. While the denotation of the noun *buildings* as such can contain only actual buildings, the denotations of *building that I could have built and buildings built in the eighteenth century* contain possible and past objects as well.[15] The reason why (18c) cannot be true must be that nouns are existence-entailing, unless they are modified by a suitable intensional modifier.[16]

---

[14]See Carlson [2] and Grosu and Landman [9].

[15]Other kinds of intensional modifiers that extend the domain of quantification of a *there*-sentence are those in the sentences below:

(i) (a) There are possible buildings that do not exist.
    (b) There are philosophers of the past who hold the same view.

[16]Past objects may allow for other predicates that do not entail existence besides psychological predicates, namely predicates describing the causal effects or historical influence of an object, such as *influential* or *important*. Even sortal predicates may in certain cases not be existence-entailing, namely in the case of individuals whose influence endures beyond their life span or whose achievements are meant to endure. Thus if A and B are two people that lived in the past, (1a) is acceptable in the present tense if A was a philosopher whose work is still known. By contrast, (1b) is not likely to be acceptable, unless B, let us say, initiated a tradition or created a lasting recipe:

1. (a) A is a philosopher.
   (b) B is a baker.

Not only *there*-sentences may range over nonexistent (past or possible) objects, also quantificational NPs by themselves can, provided they are modified by the same sorts of modifiers:

19. (a) Some buildings John could have built will never exist.
    (b) Some houses built in the eighteenth century do not exist anymore.

The semantics of intensional modifiers of existence-entailing nouns is straightforward: they act as modal operators in the definition of the property expressed by the complex nominal, as in (20):

20. $\lambda x[\Diamond \text{ building}(x)]$

The relative-clause constructions in (18a,b; 19a,b) require, as in the case of intentional verbs, that the head noun be interpreted in the lower position inside the relative clause. This permits the noun to be interpreted in the scope of the modal or temporal operator, as below:

21. (a) [that John could have built e buildings]
    (b) $\lambda x[\Diamond(\text{building}(x) \ \& \ \text{build}(\text{John}, x))\ ]$

There are various syntactic criteria for when a relative clause will involve an internal head and when not (Carlson [2]). One of those criteria is the impossibility of stacking of relative clauses. More precisely, the same type of relative clause, with the same head-internal interpretation, cannot be stacked. The example below, which does not sound very good, illustrates the constraint:

22. ?? the buildings that I could have built that could have been financed

Example (22) contrasts with the examples below, which are fine:

23. (a) the buildings that I could have built that never came into existence
    (b) the buildings that were built in the eighteenth century that do not exist anymore

The reason for the acceptability of (23a) and (23b) is straightforward. In these examples, the second relative clause, on a head-external interpretation, simply expresses a restriction on the set specified by the first relative clause. By contrast, a head-external interpretation of the second relative clause in (22) is impossible for semantic reasons, and a head-internal interpretation is unavailable for syntactic reasons: the head of the entire construction is already used for the head-internal interpretation of the first relative clause.

8.3 Intentional Verbs and Intensional Verbs

The complements of intentional verbs share some similarities with the complement of intensional transitive verbs, such as *need* and *look for*, in particular a lack of specificity and the lack of a requirement that the complement stand for an actual objects in order for the sentence to be true or false. However, intensional verbs need to be sharply distinguished from intentional verbs.[17] The complement of intentional verbs behaves just like an ordinary referential or quantificational NP, though it sets up a 'new' domain of intentional objects, objects that depend just on the intentional act described by the verb. By contrast, the complement of intensional transitive verbs, on all of the existing analyses, contributes a semantic value of a different type from that of the complement of an intentional or extensional verb. On one view, the complement contributes a quantifier (Montague [22], Moltmann [15], [19, ch. 5]), on another a property (Zimmermann [37]).

8.3.1 Nonspecificity and Underspecfication

One criterion for intensional verbs is nonspecificity. For a subclass of intensional verbs, that of verbs of absence such as *need* or *look for*, the relevant notion of nonspecificity manifests itself in the possibility of adding 'any will do' to the sentence (Moltmann [15, 17], Zimmermann [38])[18]:

24. John needs a horse, any will do.

The nonspecificity of the complement of intensional verbs needs to be sharply distinguished from the indeterminateness of intentional objects as arguments of intentional verbs. The indeterminateness of intentional objects consists in their underspecification with respect to properties, whereas the nonspecific reading of intensional verbs like *need* has to do with the semantic type of their complement, as an intensional quantifier or a property. This also means that quantificational complements such as *at least two* N display the very same nonspecific reading:

25. John needs at least two assistants.

Need is a modal verb of absence, and the semantic contribution of the complement is best understood by paraphrasing (25) as 'For any minimal situation s satisfying John's needs, there are at least two assistants John has in s' (Moltmann [15, 17], [19, ch. 5]).

There are other intensional verbs besides modal verbs of absence that take quantificational complements, but to which the 'any will do'-test will not apply. They include *owe, buy, sell, recognize,* and *find* (Moltmann [15], Zimmermann [38]).

Some intensional verbs may be intentional verbs at the same time. The psychological verb *want* is an example. *Want* can clearly take intentional objects as arguments in examples like (26), assuming that the book does not describe an actual house:

[17] The distinction is often ignored both in the philosophical and in the linguistic literature.

[18] For arguments that nonspecificity, rather than failure of substitutivity or existential quantification, is characteristic of intensional transitive verbs see Moltmann [15] and Zimmermann [38].

26. John wants the house described in the book.

But *want* also takes quantificational complements with the relevant nonspecific reading:

27. John wants at least three assistants, any will do.

Want thus is polysemous: it has an interpretation as an intensional verb, involving a property or intensional quantifier, and as an intentional verb, taking an intentional object as its argument.

8.3.2 The Choice of Neutral Proforms

Besides nonspecificity, there are two other linguistic characteristics of intensional verbs that distinguish them from extensional and intentional ones (Moltmann [15]). First, intensional verbs generally require impersonal proforms, regardless of the gender and content of the NP they replace:

28. John needs something/?? someone, namely an assistant.
29. (a) There is something/?? someone John needs, namely an assistant who speaks French.
    (b) There is something John made reference to, namely a person who speaks French fluently.

By contrast, intentional verbs generally go with proforms that match the features of the NP they replace:

30. John mentioned someone/? something, a woman (in fact, a woman that does not exist).

A related difference between the two sorts of verbs consists in the 'identity conditions' concerning what is shared by two occurrences of intensional or intentional verbs. Two distinct occurrences of intensional verbs share the same object ('the same thing') in case they would involve the same property or quantifier:

31. John needs an assistant and Mary needs an assistant, and thus John and Mary need the same thing.
32. (a) John needs the same thing as Bill, namely an assistant that speaks French.
    (b) John is looking for the same thing as Bill, a house with a garden.

By contrast, for two occurrences of intentional verbs to share the same object, they either need to share actual objects as arguments or else the intentional acts they describe need to be coordinated and thus yield the same intentional object, as illustrated by the two readings of the sentences below:

33. (a) John and Mary mentioned the same book.
    (b) John and Mary were thinking about the same woman.

The proforms that can take the position of NP-complements of transitive intensional verbs belong to a particular class of *special quantifiers*, which include *everything, nothing, the same thing*, but also the proform *that* and the relative pronoun *what*. On a common view, such quantifiers range over higher-order semantic objects, intensional quantifiers or properties (Moltmann [15], Zimmermann [37]). On an alternative view, the Nominalization Theory developed in Moltmann [16, 17], [19, ch. 5], special quantifiers are *nominalizing quantifiers* that introduce 'new' entities into the semantic structure of the sentence, entities that could also be the semantic values of corresponding nominalizations. According to that view, what is shared in (31) is 'the need for an assistant and what is shared in (32b) is 'the search for a house with a garden'.

Special quantifiers and the associated identity conditions are characteristic of intensional verbs, but not intentional verbs. The identity conditions associated with intentional verbs crucially involve the coordination of intentional acts: no two intentional objects may be 'the same N' that depend on different, uncoordinated acts. Thus, an argument such as the following is invalid, in a circumstance in which the women John and Bill mentioned do not exist and John's and Bill's acts of mentioning are not coordinated:

34. John mentioned a woman with red hair.
    Bill mentioned a woman with red hair.
    ___
    John and Bill mentioned the same woman.

The same holds for predicates describing nonlinguistic intentional acts, such as acts of imagination or (nonveridical) perception. Thus the inferences below are invalid if John's and Bill's imaginations and perceptions are not coordinated:

35. John imagined a woman with blue hair.
    Bill imagined a woman with blue hair.
    ___
    John and Bill imagined the same woman.

36. John saw a red spot.
    Bill saw a red spot.
    ___
    John saw the same spot as Bill.

If instead of *the same* N special quantifiers or pronouns occur in the conclusion, the arguments become valid, as when replacing the conclusions in (34)–(36) by (37a), (37b), and (37c) respectively[19]:

37. (a) John mentioned what Bill mentioned (a woman with red hair).

[19]The same contrast can be observed with verbs of creation, which themselves are in fact intensional verbs (Moltmann [15]):

(i) (a) John is writing a poem. Mary is writing the same thing.
    (b) John is writing a poem. ?? Mary is writing the same poem.

(ib) has only a reading on which John's literary creation miraculously coincides exactly with that of Mary, which is not something implied by (ia).

(b) John imagined the same thing as Bill (a woman with blue hair).
(c) John saw the same thing as Bill (a red spot).

But what (37a,b,c) report is the sharing of a type of object, not a single intentional object. The reason why special quantifiers and pronouns are tolerated in (37) is that the intentional verbs are interpreted as intensional verbs. Such a reinterpretation is available for at least some intentional verbs, by a form of 'type-shifting'. Just as Montague Grammar allows extensional verbs to be 'type-shifted' to intensional verbs, intentional verbs can be 'type-shifted' to intensional ones, as roughly below (adopting, for the sake of simplicity, the view that intensional verbs take properties as arguments):

38. For an intentional verb V, an event e, an object x, a property P, and an intentional object y, V_{intens}(e, x, P) iff V(e, x, y), where for some event e' coordinated with e, e' attributes P to y.

The same difference in identity conditions displayed by *the same* N and *the same thing* are displayed by two types of anaphoric pronouns. Anaphoric *that* classifies with *the same thing* whereas *it* classifies with *the same* N. *That* cares about the identity of a type, whereas *it* requires the identity of an intentional object. This is why (39a) and (40a) below are acceptable, whereas (39b) and (40b) are not, on a natural reading on which the mental states described are not coordinated:

39. (a) John imagined a circle. Mary imagined that too.
(b) John imagined a circle. ?? Mary imagined it too.
40. (a) John wants a nice book. Mary wants that too.
(b) John wants a nice book. ?? Mary wants it too.

The second sentence in (40b) is unacceptable unless there is a particular book, existent or intentional, that both John and Mary want.

8.3.3 The Semantics of Special Quantifiers with Intensional Verbs

On the 'Nominalization Theory', special quantifiers have a 'nominalizing' function, ranging over entities that would be semantic values of a corresponding nominalization. Thus, *what Mary needs* below would stand for 'the need for a house', which is said to be something John shares with Mary:

41. John needs what Mary needs, a house

Part of the motivation for the Nominalization Theory comes from the predicates that can apply to special quantifiers. Such predicates generally cannot be understood as predicates of higher-order semantic objects such as intensional quantifiers or properties. For example, in (42a,b) *count* and *unusual* can hardly be understood as predicates of semantic objects such as quantifiers or properties:

42. (a) John counted what he needs.
    (b) John needs something unusual.

Rather *count* and *unusual* in (42a,b) apply to entities of the sort of John's needs.

Whereas (42a,b) involves quantification over particular entities of the sort 'John's needs', (41) involves quantification over kinds, entities of the sort 'the need for a house' (which has both 'John's need for a house' and 'Mary's need for a house' as instances).

There are cases which at first sight present problems for the Nominalization Theory, namely acceptable examples such as those below (Moltmann [17], [19, ch. 5]):

43. Mary has what she needs, a house.

Such cases however, are not a counterexamples to the Nominalization Theory, but simply require a modification of it. What is special about (43) is that the situation described in the main clause, Mary owning a house, is a situation satisfying her need. *What she needs* in (43) does not stand for Mary's need, but rather for the satisfier of Mary's need's, or rather a variable satisfier of her need. (Moltmann [19, ch. 5]). A variable satisfier of a need is a variable object that has manifestations as ordinary objects in different circumstances and may lack a manifestation in the actual circumstances. In general, special quantifiers and pronouns with transitive intensional verbs stand for such variable objects, rather than what the nominalizations of the verbs would stand for, entities like needs.

8.3.4 Relative Clause Constructions with Intensional and with Intentional Verbs

NPs formed with relative clauses with intensional verbs and with intentional verbs form referential NPs, NPs that are arguments of ordinary predicates and that can act as antecedents of anaphora, as below:

44. (a) The woman John described is American. She is fluent in French and German.
    (b) The assistant John needs must speak French. He should also be fluent in English.

Whether or not the woman John described exists does not bear on the acceptability of (44a), and John's not having an actual assistant does not bear on the acceptability of (44b). NPs modified by relative clauses with intentional verbs and with intensional verbs differ, though, in what they stand for. There are two semantic reflections of that.

First of all, it is reflected in the applicability of the predicate *exist*. NPs with intentional verbs can stand for intentional objects of which *exist* is false, as in (45a), but those with intensional verbs cannot, as seen in (45b):

45. (a) There is a book John mentioned that does not exist.
    (b) ?? There is an assistant John needs that does not exist.

(45b) is acceptable only on a reading on which the indefinite characterizes a type of object, rather than a particular actual or intentional object—unless of course the verb has in fact the extensional rather than the intensional reading. (45b) contrast with (45c) with a psychological verb of absence, which better tolerates *exist*:

(c) There is a book John wants that does not exist.

Similarly, (46b) is not really worse than (46a):

46. (a) There is a woman John is thinking about that does not exist.
    (b) There is a woman John is looking for that does not exist.

The reason is that psychological verbs generally have a variant as intentional verbs, leading to a domain of intentional objects.

There is another semantic difference between NPs with intentional verbs and with intensional verbs. Definite NPs modified by a relative clause with an intensional verb generally are subject to the Modal Compatibility Requirement, the obligatory presence of a modal in the main clause (Moltmann [18, 21])[20]:

47. (a) The assistant John needs must speak/may speak/??? speaks English.
    (b) The woman John is looking for must be/may be/??? is tall and blond.

By contrast, definite descriptions with a relative clause containing an intentional verb are subject to no such requirement:

48. (a) The woman John is dreaming about is tall and blond.
    (b) The building John described is made almost entirely of glass.

While definite NPs modified by relative clauses with intentional verbs describe intentional objects, definite NPs with intensional verbs describe *variable objects* of a certain sort (Moltmann [18], [19, ch. 5], [21]). Variable objects are entities that have different manifestations as ordinary objects in different circumstance and may lack a manifestation in the actual circumstance. For example, 'the president of the US' viewed as a variable object will be an entity that has as its manifestation at a time and a world whoever is president of the US at that time in that world. 'The assistant John needs' viewed as a variable object is a *variable satisfier* of 'John's need', which means it is a variable object that has manifestations in exactly those situations (exactly) satisfying John's need and that has as its manifestation in a situation exactly satisfying John's need the assistant that John 'has' in that situation (which means that John stands in a contextually relevant relation *R* to):

[20]The Modal Compatibility Requirement has been noted first for the related construction *the gifted mathematician John claims to be* by Grosu and Krifka [8]. See also Moltmann [21].

49. [*the assistant John needs e*] = the variable object o such that o has manifestations in exactly those situations s such that for some e, need(e, John), s (exactly) satisfies e, and for some e, need(e, John), for any situation s satisfying e, manif(o, s) = ιx[assistant$_s$(x) & R$_s$(x, John)], for a relevant contextually given relation R.

The modal is required because it allows accessing nonactual manifestations of the variable object for the purpose of applying the predicate, by the condition below (Moltmann [18]):

50. A variable object has a (circumstance-relative) property P in a circumstance s iff if the manifestation of o in s has P in s (P$_s$(manif(o, s)))

(50) is a general condition on applying a property to variable objects.

8.4 Actual Objects Acting as Intentional Objects?

The present account sharply distinguishes between intentional objects and actual objects as denotations of the complements of intentional verbs. There are examples, however, where actual objects appear to qualify as intentional objects:

51. (a) John now lives in the house he had dreamt of.
    (b) John finally bought the house he had always longed for.
    (c) John now owns the car he had always been fantasizing about.

(51a) suggests that the actual house John lives in qualifies as the house John dreamt of, which means that the latter is in fact an actual object (though of course, John did not dream of all the aspects of the house he now lives in). This would mean that at the time of his dreaming John's dream is directed toward an actual object not an intentional one, unbeknownst to him. But this is not plausible: a different house than the one John lives in could have fulfilled John's dreams just as well. There are constraints on when actual objects may qualify as the objects of thought or imagination: there needs to be a causal connection to the intentional act or state; having certain properties is not enough. The examples in (51) in fact turn out to be special. In many cases, an actual object meeting the conditions on the object of thought does not make the construction in question acceptable:

52. (a) ?? Yesterday John saw the castle he had imagined.
    (b) ?? John noticed the car he had always been fantasizing about.
    (c) ?? John now lives in the house Mary once thought about.

Problematic are also cases where there are in fact several actual objects that meet the conditions on the object of thought, for example (53), in a context in which John dreamt of a 'generic' castle and then saw several that match the one he had dreamt of:

53. ?? John saw several castles he had dreamt of.

The reason for the acceptability of (51a–c) thus cannot be an actual object meeting the conditions on an object of thought. In (51a), the main clause describes an actual situation that satisfies John's dream, which is not just directed toward an object: John's dream is not just a dream about a house, but a dream about a house to live in. On the shared reading, the complement of *dream* does not in fact describe an intentional object, but rather it acts as the complement of an intensional verb, specifying a variable satisfier of the event of dreaming. That is, *the house John had dreamt of* stands for a variable object that in each situation satisfying John's dreams has a manifestation that is a house John lives in. Intensional verbs of absence can share their 'object' with extensional verbs in case the intensional verb describes a situation that is a satisfaction situation for the intensional verb as in the case of (43) repeated below:

54. Mary has what she needs, namely a house.

In such case a case, more precisely, the extensional and the intensional share a variable satisfier. This means that the acceptability of (51a–c) is due to an intensional interpretation or rather reinterpretation of the intentional verb.

8.5 Generic and Intentional Objects

There are certain predicates that take objects similar to intentional objects, but that in fact should be viewed as generic objects rather than intentional objects. These are predicates of comparison, which include verbs of resemblance as well as comparatives.

The arguments of those predicates are generally given by indefinites:

55. (a) This animal resembles a unicorn.

Predicates of comparison, on the intensional reading, do not permit quantificational complements:

(b) This animal resembles at least one unicorn.

Predicates of comparison differ in that respect from intensional verbs of absence, such as *need*, which allow for quantificational complements.

Predicates of comparison may also take as arguments intentional objects, introduced by intentional verbs:

56. (a) Mary resembles the woman John talked about.
    (b) Mary is more intelligent than the woman the novel describes.
    (c) The building is taller than the building John had described.

In fact, intentional objects can fill the two argument positions of a comparative, just like generic objects[21]:

57. (a) The woman John described is more attractive than the woman Bill described.
    (b) A unicorn is smaller than a dragon.

The indefinite complements of predicates of comparison do not seem to have a predicative function. More plausibly, they stand for generic objects, just as plausibly in (58):

58. (a) the example of a unicorn
    (b) the description of a unicorn

Like intentional objects, there is nothing more to generic objects than their partial qualitative nature. Unlike intentional objects, however, generic objects are intuitively not 'nonexistent' objects. Rather they appear to be neutral as to existence or nonexistence: the existence predicate seems neither true nor false of them:

59. (a) ?? Charlie resembles a unicorn, which does not exist.
    (b) ?? Charlie resembles a pony, which exists.
    (c) Charlie resembles the man Bill described who does not exist.

Instead of being existent or nonexistent, generic objects are instantiated, exemplified, or manifested in particular objects. Comparative predicates, it appears, are able to apply to entities on the basis of their qualitative specification only, regardless of their existential status.

It is interesting to note that intentional verbs can also take generic objects, as on a generic reading of the verbal complement in (60a,b):

60. (a) When asked about examples of mythical beasts, John mentioned a unicorn.
    (b) John described a unicorn.

Intentional verbs such as *imagine* naturally take generic objects as arguments, allowing for co-predication with a predicate of comparison:

61. Charlie resembles the animal John imagined.

Such generic objects need to be sharply distinguished from intentional objects. If an existence predicate is applied to a generic object, it can state only the existence of a particular instance, not the existence of the generic object as such:

---

[21]The acceptability of (57b) is quite surprising in fact, since the positive would not allow for a singular generic indefinite, as is familiar from the linguistic literature on generics:

(i) ?? A unicorn is small.

The difference obtains whether or not the NPs range over existent or nonexistent objects. Thus, it also obtains for the examples below:

(ii) (a) A mouse is smaller than an elephant.
     (b) ?? A mouse is small.

I do not know of an explanation of this difference.

62. The animal John imagined exists: there are ponies.

Predicates of comparison take generic objects for just the same reason, it appears, that they take intentional objects: they care about the properties of objects only, not whether or not the objects exist.

8.6 Negative Existentials

Negative existentials with descriptions of intentional objects based on intentional verbs have hardly received attention in the philosophical literature. The more familiar negative existentials involve a proper name or ordinary definite description in subject position:

63. (a) The king of France does not exist.
    (b) Vulcan does not exist.

Such negative existentials arguably involve intentional objects as well. In fact, McGinn [11] argued that apparently empty terms in negative existentials stand for intentional, nonexistent objects, in roughly the present sense, as entity constituted by failed intentionality. Obviously in that context, intentional objects would not depend on a described intentional act or state. Rather, they would depend on a contextually given quasi-referential act. Thus in (63a), the subject presumably relates to a pretend act of reference, or better a simulated unsuccessful act of reference, by a recent or contemporary philosopher; in (63b) the subject presumably relates to an attempted act of reference on the part of astrophysicists in the past. More precisely, the utterance of the subject will be a pretend act of reference coordinated with a contextually given quasi-referential act involving the same name or description.

There is a range of evidence (not considered by McGinn) that the subject of a negative existential stands for an intentional object, an object dependent on a quasi-referential act. First, not any ordinary definite description is acceptable as the subject of a negative existential. The definite descriptions in the following examples are appropriate only insofar as their use is coordinated with a relevant previous quasi-referential use of the same description[22]:

64. (a) Mary's child does not exist.
    (b) The tree in the garden does not exist.

(64a) cannot just be used to state that Mary does not have a child and (64b) cannot be used to state that the garden does not have a tree. Rather someone must have tried to refer to Mary's child or the tree in the garden before.

[22]The philosophical literature also discusses the following sentence:

(i) The largest natural number does not exist.

This sentence seems to me to be subject to the same condition involving a previous quasi-referential act, pace Russell's [25] account of definite descriptions acting as quantifiers in such sentences.

Furthermore, not just any name that fails to have a referent can act as the subject of a negative existential. Newly created names cannot act in that function, and neither can names whose use does not relate to a preceding quasi-referential act. For example, an expression that could be a proper name, but has not been used as such, let's say *Barkab*, cannot felicitously occur in the subject position of a negative existential:

65. ??? Barkab does not exist.

Such intuitions are unaccounted for on a view on which the subject of negative existentials such as (64a,b) is empty and negation is understood as external negation (Salmon [27, 28]), unless it is supplemented by conditions on previous name-using practices (Sainsbury [26]) (this, though, would not carry over to definite descriptions).

The view that the subject of a true negative existential stands for an intentional object also accounts for the intuition that with non-referring singular terms negative existentials are false rather than just not true: *the golden mountain exists* is simply false, rather than neither true nor false, unlike a sentence such as *the present king of France is bald*, where there is a good intuition that the sentence is truth-valueless.

In addition, Salmon's view that negation in negative existentials is external negation is problematic. On that view, *not* in a negative existential would be the same kind of negation as in (66), which is naturally followed by a *because*-clause:

66. The king of France is not bald, because there is no king of France.

But (66) involves a particular intonation, namely a focus on *not*, rather than, as with ordinary negation, the predicate. By contrast, in a sentence with *exist* it is the predicate that is focused. That is, negative existentials do not appear to be cases of 'metalinguistic negation' in the sense of Horn [10]. Another problem for the view that negation in negative existentials is external is that negation as 'external negation' should be negation taking wide scope over the subject. However, with a quantificational subject, no wide-scope can be attested, unless *not* is strongly focused:

67. (a) Every one we talked about does not exist.
    (b) At least two people we talked about do not exist.

The treatment of negation as external negation also has difficulties with the sentence below:

68. Every one we talked about except Anna Karenina exists.

Except also involves negation, but negation here could hardly end up as external negation in the logical form of the sentence.

Thus, the subject of a negative existential does not appear to have a special semantics, involving an empty denotation that triggers an external interpretation of negation. Rather it is on a par semantically with referential NPs, and negation in negative existentials is just ordinary negation.[23]

Turning to the semantics of negative existentials, the entire subject NP will be coindexed with a quasi-referential act given by the linguistic or nonlinguistic context. For proper names, I will, for the present purposes, make the same assumption as for common nouns when used to possibly describe intentional objects: proper names have a two-place variant with an additional event argument place for an act of attribution. Simplifying, the interpretation of a proper name coindexed with a contextually given quasi-referential act will then be as below:

69. (a) $[\textit{Anna Karenina}_i]^{e_i} = \iota x[x = f(e_i)\ \&\ \textit{Anna Karenina}(x, e_i)]$

That is, the referent of the name *Anna Karenina*, co-indexed with some event *e* is the intentional object projected from *e* that involves the attribution of the name in an act coordinated with *e*.

Similarly, a definite description coindexed with a quasi-referential act given by the context will be interpreted as follows:

(b) $[\textit{the king of France}_i]^{ei} = \iota x[x = f(e_i)\ \&\ \textit{king of France}(x, e_i)]$

That is, the referent of *the king of France* coindexed with an intentional event *e* stands for the intentional object dependent on *e*, which is coordinated with an act of the attribution of the property of being king of France.

One question that arises with negative existentials is, why could the subject not stand for a fictional character, so that (63a,b) would in fact come out false? Fictional characters obviously can be referents of terms making explicit reference to them such as *the fictional character Anna Karenina*. The question then is, under what circumstances can 'nonreferring' names and descriptions stand for fictional characters, rather than intentional objects? I will restrict myself to just a few observations and generalizations. First of all, there clearly are contexts in which empty names and definite descriptions can stand for fictional characters rather than intentional objects[24]:

70. (a) Anna Karenina was created by Tolstoy.
    (b) Anna Karenina serves as a model for an unhappily married, intelligent woman.

[23] Another option one might think of would be to take *the king of France* to stand for a merely possible object and say that it does not exist. But see Kripke [5] for a critique of that view.

[24] Another apparent case of reference to a fictional character is (i) below:

(i) Anna Karenina is an interesting fictional character.

However, *fictional character* has a 'reifying' function in this context, mapping a presentation of a name (a non-referential use of the name) onto a fictional character of which *interesting* is then predicated. It is the same function that *fictional character* has in *the fictional character Anna Karenina*, where it guarantees reference to a fictional character on the basis of a non-referential use of the name *Anna Karenina*. See Moltmann [19, ch. 6], for a discussion of the reifying function of certain sortals in predicate position.

In negative existentials, as in (63a, b) and below, an empty name or definite description cannot refer to a fictional character:

71. Anna Karenina does not exist.

It is impossible to get a reading of (71a) on which the sentence come out false.[25] This is in contrast to a negative existential with a subject making explicit reference to a fictional character:

72. The fictional character Anna Karenina does not exist.

The reason why the subject in (71) and (63a, b) *has* to stand for an intentional object and cannot stand for fictional characters may be the following. A term can stand for an intentional object in only those contexts in which it could have been used to refer in the ordinary way (as in the case of *exist*), and it can be used to refer to a fictional character only in those contexts that exclude ordinary reference (as in the case of the verb *create*).

8.7 Restrictions on Predication of Intentional Objects

Only certain predicates, we have seen, can be predicated of intentional objects, such as intentional verbs and (negated) existence predicates, as well as predicates of evaluation and comparison. To this generalization the observation needs to be added that when a subject stands for an intentional object, then any sort of predicate can follow:

73. (a) The person John described is a woman who knows many people.
    (b) The woman the book is about is someone that likes everyone.

The restriction to the subjects is crucial, though, for applying an ordinary predicate. The sentences below are impossible if there is no actual woman the book is about:

74. (a) Many people know the woman the book is about.
    (b) Everyone met the woman the book is about.

The restriction to subjects indicates that for ordinary properties to be predicated of intentional objects, the intentional act on which the intentional object depends needs to be accessible for the predicate to be predicated of the intentional object. Given standard syntactic views, the intentional act involved in the interpretation of the subject is accessible for the predicate (the subject c-commands the predicate), but not so for the intentional act involved in the interpretation of the object (the

[25] See Thomasson [30]. In some of the literature, the intuition is not quite recognized as such, for example in Salmon [27].

object does not c-command the predicate). Let us first take the case of proper names in subject position as in (75a):

75. (a) Anna Karenina is a woman.
 (b) [Anna Karenina]^i is [a woman$_i$]^i
 (c) woman(ιx[x = f(e$_i$) & *Anna Karenina*(x, e$_i$)], e$_i$)

The syntactic configuration of (75a) allows sharing of an index of the subject and the predicate, as in (75b), which can be interpreted as the relativization of the interpretation of the subject and the predicate to the same contextually given intentional act as roughly in (75c).

Definite descriptions with intentional verbs as in (73a,b) are more difficult to handle. On the analysis employed so far, the intentional act on which the intentional object depends is introduced by an existential quantifier inside the definite description operator. In order for that event to be accessible for the interpretation of the predicate, the existential quantifier would have to be understood dynamically, allowing it to bind variables outside its scope. Note that an ordinary predicate when predicated of an intentional object may at the same time combine with a negative existential:

76. The person the book is about is a woman that does not exist.

This means that not all of the predicate will be interpreted relative to the intentional act on which the subject depends, namely not the relative clause.

8.8 Conclusions

This paper has argued that transitive intentional verbs go along with a special semantic interpretation of their complement, involving intentional objects in a particular, event-dependent sense. This semantics is quite different from the semantics of the complement of transitive intensional verbs, which has long been recognized as being special. The difference in semantic interpretation of constructions with intentional and with intensional verbs accounts for a range of linguistic differences, in particular with quantificational complements.

Intentional objects, I have argued, do play a role in the semantic structure of natural language and are reflected in particular linguistic constructions. But the role of intentional objects is strictly limited: intentional objects in general can be made available only by the presence, in the semantic structure of the sentence, of intentional acts on which the intentional objects depend. This view of the involvement of intentional objects in the semantics of natural language is thus significantly different from the standard Meinongian view, whose ontology is much harder to accept.

References

1. Brentano, F. 1874/1911. *Psychologie vom Empirischen Standpunkt*. Leipzig: Duncker & Humblot. (English trans. 1995. London: Routledge).
2. Carlson, G. 1977. Amount relatives. *Language* 53: 520–542.
3. Crane, T. 2001. Intentional objects. *Ratio* 14: 336–349.
4. Everett, A. 2000. Referentialism and empty names. In [6].
5. Kripke, S. 2013. Reference and existence. the John Locke lectures. New York: Oxford UP.
6. Everett, A. and T. Hofweber (eds.). 2000. *Empty names, fiction, and the puzzle of non-existence*. Stanford: CSLI Publications.
7. Fine, K. 2007. *Semantic relationalism*. Oxford: Oxford University Press.
8. Grosu, A., and M. Krifka. 2007. The gifted mathematician that you claim to be. *Linguistics and Philosophy* 30: 445–485.
9. Grosu, A., and F. Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6: 125–170.
10. Horn, L. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61: 121–174.
11. McGinn, C. 2000. *Logical properties*. Oxford: Oxford University Press.
12. Meinong, A. 1904. Gegenstandstheorie. In Untersuchungen zur Gegenstandstheorie und Psychologie, ed. A. Meinong. Leipzig. (English trans.: The theory of objects). In *Realism and the background to phenomenology*, ed. R. Chisholm. London: Allen & Unwin.
13. Miller, B. 1975. In defense of the predicate 'Exist'. *Mind* 84: 338–354.
14. Miller, B. 1986. 'Exists' and existence. *The Review of Metaphysics* 40: 237–270.
15. Moltmann, F. 1997. Intensional verbs and quantifiers. *Natural Language Semantics* 5(1): 1–52.
16. Moltmann, F. 2003. Nominalizing quantifiers. *Journal of Philosophical Logic* 32: 445–448.
17. Moltmann, F. 2008. Intensional verbs and their intentional objects. *Natural Language Semantics* 16(3): 257–281.
18. Moltmann, F. 2012. Intensional relative clauses and the notion of a variable object. In *Proceedings of the 18th Amsterdam colloquium 2011*, Amsterdam. FOLLI lecture notes in computer science (LNCS), 431–440. Springer.
19. Moltmann, F. 2013. *Abstract objects and the semantics of natural language*. Oxford: Oxford University Press.
20. Moltmann, F. 2013. The semantics of existence. *Linguistics and Philosophy* 36(1): 31–63.
21. Moltmann, F. (2015). Variable objects and truthmaking. In *The philosophy of Kit Fine*, ed. M. Dumitru. Oxford: Oxford University Press.
22. Montague, R. 1973. The proper treatment of quantification in ordinary English. In *Approaches to natural language*, ed. J. Hintikka, et al., 221–242. Dordrecht: Reidel. (Reprinted in Thomason, R. 1974. Formal philosophy. Selected papers by Richard Montague. Montague Grammar.).
23. Parsons, T. 1980. *Nonexistent objects*. New Haven: Yale University Press.
24. Priest, G. 2005. *Towards non-being: The logic and metaphysics of intentionality*. Oxford: Oxford University Press.
25. Russell, B. 1968. On denoting. In: *Logic and knowledge*, ed. R.C. Marsh. London: Allen & Unwin.
26. Sainsbury, M. 2005. *Reference without referents*. Oxford: Oxford University Press.
27. Salmon, N. 1987. Existence. *Philosophical Perspectives* 1: 49–108.
28. Salmon, N. 1998. Nonexistence. *Nous* 32(3): 277–319.
29. Taylor, K. 2000. Emptiness without compromise. In [6].
30. Thomasson, A. 1999. *Fiction and metaphysics*. Cambridge: Cambridge University Press.
31. Tye, M. 1984. The adverbial theory of visual experience. *Philosophical Review* 93: 195–225.
32. van Inwagen, P. 2000. Quantification and fictional discourse. In [6].
33. van Inwagen, P. 2008. McGinn on existence. *The Philosophical Quarterly* 58: 36–58.
34. Voltolini, A. 2009. Consequences of schematism. *Phenomenology and the Cognitive Sciences* 8(1): 135–150.

35. Voltolini, A. 2009. Seven consequences of creationism. *Metaphysica* 10: 27–48.
36. Walton, K.L. 2000. Existence as metaphor. In *Empty names, fiction and the puzzles of nonexistence*, ed. A. Everett and T. Hofweber, 69–99. Stanford: Stanford University Press.
37. Zimmermann, T.E. 1993. On the proper treatment of opacity in certain verbs. *Natural Language Semantics* 1: 149–179.
38. Zimmermann, T.E. 2001. Unspecificity and intensionality. In *Audiatur Vox Sapientiae*, ed. C. Féry and W. Sternefeld, 514–532. Berlin: Akademie.

Part III
The Carnap-Quine Legacy

Chapter 9
Life on the Range: Quine's Thesis and Semantic Indeterminacy

G. Aldo Antonelli

Abstract This paper explores the ramifications of Quine's Thesis that ontological commitment is determined on the basis of the range of quantified variables in light of a non-standard semantics for the first-order quantifiers analogous to Henkin's general interpretations for second-order logic. On such a general semantics, the specification of the meaning of the first-order quantifiers requires, besides the selection of a first-order domain D of objects (as is customary), also the further identification of a second-order domain of subsets of D. In the absence of such a further identification, the semantics of "there is" and "for all" is radically indeterminate. Moreover, the general semantics might well be "transparent" in that no *semantic* facts might be available to discriminate the general from the standard interpretation. The general semantics is rich in consequences for Quine's thesis: the paper shows how the two halves of the thesis, i.e., the symmetric claims that being the value of a bound variable is necessary or, respectively, sufficient for ontological commitment fail, albeit in interestingly different ways. The result undermines the prospects of philosophical ontology construed as the quintessentially armchair project of extracting ontological commitments from the semantic analysis of quantified statements.

9.1 Introduction

Philosophers of all stripes have, with little or no hesitation, entrusted the first-order quantifiers "there is" and "for all" with an extraordinary task—that of carrying the ontological commitments of theories, be they the informal theories implicit in everyday reasoning, the formal or semi-formal theories of science, or the lucubrations of metaphysicians. The practice rests on a popular simplification of

G.A. Antonelli (✉)
Philosophy Department, UC Davis, 2296 Social Sciences and Humanities Building, Davis, CA 95616, USA
e-mail: antonelli@ucdavis.edu

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_9

Quine's thesis that "to be is to be the value of a variable" [25, pp. 34–35], which has come to be regarded, at least by some, as the basis for identifying the class of entities whose existence is necessary for the truth of a given body of propositions. Quine is, in fact, rather more careful in his characterization than contemporary philosophical practice takes him to be: not only does he specify that only *bound* variables are relevant here, as is obvious, but also that the criterion itself cannot be used to adjudicate between different ontologies, since it can only be used in "testing the conformity of a given ... doctrine to a prior ontological standard." This will be our conclusion as well, although by a different route.

Upon reflection, the Quinean thesis turns out to be remarkably more mysterious than it might appear at first. What, in fact, could it possibly mean to be *the value* (or even *a value*) of a bound variable? Strictly speaking, bound variables have *no* values: a point that was already recognized by Russell, when he pointed out that there is no (specific or—assuming the notion makes sense—generic) individual of which $\exists xPx$ predicates the property P. Quantified variables are purely *syncategorematic* devices, as the scholastics well knew, and too many of the moderns seem to have forgotten (see [6]). We can agree, though, that for an object a to be the value of a bound variable it means for a to satisfy a formula $\varphi(x)$, thereby functioning as a truth-maker for the corresponding existentially quantified sentence $\exists x\varphi(x)$. The other side of the Quinean equation is similarly obscure: it is not clear precisely what it means for a given theory to be ontologically committed to X. It turns out that explicating this notion is a somewhat subtler task than one might anticipate (see, for instance, [21], or [27] for a more recent treatment): in what follows we will assume that an intuitive understanding of what ontological commitment amounts to is available, and that will be enough for our purposes. Further, a case can be made that in its original form the Quinean thesis was meant to apply to languages that have been appropriately regimented in their form and interpretation. While this fact has often been neglected by the proponents of Quine's thesis, the points we are going to make apply equally well, in fact perhaps more clearly, to the regimented languages Quine had in mind.

Quine's characterization of ontological commitment as flowing from the semantics of the quantifiers is at the basis of his suspicion of second- and higher-order logic. For if higher-order quantifiers, like their first-order counterparts, also enact ontological commitments, then it seems there is no avoiding embracing a universe replete with predicates, relations, predicates of relations, and so on. Since the least ontologically extravagant way to implement such a universe is in set-theoretic terms, Quine's [26] characterization of second-order logic as "set theory in sheep's clothing," readily follows: for second-order quantifiers rest on the set-theoretic notion of the power set, a notion that is quintessentially mathematical.

Much armchair philosophizing has ensued from Quine's proposed correlation of ontology with the semantics of first-order quantification. This is perhaps most evident in various branches of the philosophy of mathematics and mathematical philosophy: ascriptions of number, as in "the number of the planet is eight," for instance, not only employ numerical terms that purport to refer to a particular kind of abstract objects, *viz.*, numbers, but employ them in such a way that makes them available for existential generalization, allowing the inference to "something numbers the planets." But other branches of philosophy are far from immune from deploying such arguments: debates about the ultimate constituents of reality, which one would expect to be conducted with a modicum of input from science (as argued for instance by [32]), are instead at least sometimes cast as questions about first-order quantification—as though the answer to van Inwagen's [33] famous question of whether there are tables, or just simples arranged table-wise, depended on the semantics of "there is" and "for all." In what follows we will distinguish the two halves of the Quinean thesis, *viz.*, the sufficiency thesis (being the value of a bound value is a sufficient condition for ontological commitment) and the converse necessity thesis (being such a value is a necessary condition of ontological commitment). The two halves, we will see, are importantly different.

Now, like all orthodoxies, the equivalence (if not the identification) of ontological commitment and first-order quantification has spawned more than its share of heretics, too (we will consider some of the arguments in Sect. 9.2). But the heretics have, in overwhelming majority, questioned the implications of Quine's thesis for the status of second- (and higher-) order logic, whereas it has come to be accepted as a matter of course that, in the *first-order* case, the semantics of the quantifiers really is a reliable guide to ontology. Accordingly, the main heretic arguments have, in one way or another, supported the conclusion that second-order quantifiers are really no more ontologically committed than the first-order ones. We will argue that the heretics are right that, in important ways, the first- and the second-order quantifiers are very much alike, but they have the direction of the analogy backwards. Our main thesis will be that, on a deeper understanding of the nature of the quantifiers as predicates of predicates, first-order quantifiers are just as dependent on the prior specification of a second-order domain. And while, as first pointed out by Henkin [10], such second-order domains can vary in non-standard ways, likewise first-order quantifiers also admit of non-standard interpretations, a fact that has immediate repercussions for the Quinean thesis.

The development of the basic ideas underpinning such non-standard, or "general" interpretations will occupy the central parts of this paper. But it is important now to appreciate that just like the *semantic indeterminacy* of second-order quantifiers can be traced back to the existence of non-standard interpretations, the same is true in the first-order case. It has been known since Henkin's pioneering work that fixing the meaning of a second-order quantifier requires the selection, along with a first-order domain D_1 of objects, also of a second-order domain D_2 comprised of subsets of D_1, over which the quantifier is taken to range. Perhaps the central point of the present paper is that specification of the meaning of a *first-order* quantifier also requires—somewhat unexpectedly—the selection of a second-order domain, *even when* the first-order domain D_1 is already fixed. As a consequence, we will argue, first-order quantifiers are as semantically indeterminate as their second-order counterparts. Moreover, and perhaps just as importantly, such general or non-standard interpretations cannot be dismissed offhand, for there is a crucial sense in which non-standard interpretations are *transparent* from the point of view of language users, in that standard and non-standard interpretations cannot be set apart by consideration of which sentences are true or false in each.

The claim that first-order quantifiers are semantically indeterminate is of course not new. For instance, [11] first called attention to such indeterminacy under the rubric of “quantifier variance.” Although [11] and his interlocutors are careful not to characterize quantifier variance in terms of varying domains, it is difficult to give a formally precise account of such variance in terms that, in some form of other, do not presuppose different domains of quantification (but see Sider [30], for instance, for such an attempt, and Rossberg (2011, On the logic of quantifier variance, Unpublished ms) for some criticisms). This is not an issue that will occupy us in what follows, except to point out that even if quantifier variance is construed as mere variance in the domain of quantification, indeterminacy of first-order quantifiers persists *unless and until* an appropriate second-order domain is specified along with the first-order one, a feat that would appear to exceed the resources required for competent use of quantified expressions.

As we develop our account of the semantic indeterminacy of first-order quantifiers and its consequences for Quine’s thesis, an issue that will play a role is the question of why the two quantifiers “there is” and “for all” (rather than others) occupy such a central position. The question is relevant from both the orthodox and the heretic point of view: an answer is crucial from the orthodox point of view if the two standard quantifiers are to carry the full weight of our ontological commitments; but an answer is due also from the heretic point of view, for it is of limited use to argue against the ontologically loaded nature of the standard quantifiers “there is” and “for all,” if other, perhaps more fundamental quantifiers are standing at the ready, waiting to step in to do the grunt work of carrying ontological commitments.

9.2 The Quinean Critics

The Quinean thesis establishing a tight connection between the semantics of the quantifiers and ontology has been variously challenged. Some of these challenges are particularly relevant for the account to be developed below, in that they highlight features of higher-order quantification that will turn out to be equally applicable at the first order level.

Arthur Prior was perhaps the first to point out, in characteristic iconoclastic fashion, that Quine’s thesis “is just a piece of unsupported dogma” [24, p. 48]. This conclusion occurs as part of an extensive defense of *non-nominal quantification*—the idea that expressions of any syntactic category, not just names, are available for the purpose of instantiating quantifiers of the corresponding kind. In fact, non-nominal quantification is ubiquitous in natural language. Consider the sentences:

I hurt him somehow.
He’s something I am not — kind.

These sentences are obtained by existentially quantifying out the adverbial phrases in the following:

I hurt him by treading on his toes.
He is kind, but I am not.

But this fact does not by itself commit us to the existence of "ways" in which one can hurt or abstract entities such as attributes (see for a similar point [28]). In other words—to steal yet another efficacious slogan from [26]—this is an instance of "logic in wolf's clothing."

What Prior's point really comes down to, is the idea that if ontological commitment is preserved by anything, it must be preserved by converse logical entailment. In other words,

Prior's insight: If φ follows from ψ then φ's ontological commitments are also ψ's ontological commitments.

This is a principle that is supported by any conception that makes preservation of truth on any interpretation a necessary (and perhaps also sufficient) condition for logical consequence: if ψ entails φ, then φ is true on any interpretation on which ψ is true, so that the truth of φ cannot require (although of course it may allow) the existence of any further entities not already required by the truth of ψ. Hence, φ's ontological commitment is no more extensive than that of ψ. When applied to higher-order quantification, the argument neatly delivers that existential quantification over predicates carries along no more extensive ontological commitment than the sentence from which it follows. In particular, since Pa validly entails $\exists X\, Xa$, where P is an atomic predicate symbol, the second-order existential $\exists X$ is, on this account, ontologically neutral. (The case is not as clear-cut for the entailment from $\Phi(a)$ to $\exists X\, Xa$, where $\Phi(a)$ is a complex predicate expression, for then the validity of the argument depends on precisely which instances of the second-order comprehension axiom are available. But the atomic case is enough to establish Prior's point.) A version of the preservation of ontological commitment by converse entailment is also endorsed by Linnebo [17, 18], although [18] runs the implication in reverse, to conclude that indeed $P(a)$ is already committed to second-order entities (since its existential generalization is likewise committed).

A similar assessment of the tenuous connection between quantification and ontology is forcefully made in [34, p. 153]:

Neutrality: Quantification into the position occupied by a particular type of syntactic constituent in a statement of a particular form cannot generate ontological commitment to a kind of item not *already* semantically associated with the occurrence of that type of constituent in a true statement of that form.

The Neutrality Thesis embodies an important point, suggesting a partial decoupling of quantification and ontological commitment ([4] provides perhaps the most far-reaching argument for severing the connection of semantics and ontology). The part that is interesting for our purposes is the idea that to the extent to which quantification of any kind is connected to ontological commitment, such a connection is grounded on a prior selection of a domain of quantification. Thus we have a reversal in the order of dependence: rather than *enacting* such a commitment, quantification to a large extent *presupposes* it.

However, the main idea behind such reversal, important as it might be, is still not deployed in its full generality either in Prior's Insight or through the Neutrality Thesis, both of which share the narrow focus on the second-order case. Prior's argument, in its full generality, applies just as well at the first order as it does at the second order: since Pa validly entails $\exists xPx$, the ontological commitments of the latter do not exceed those of the former. And similarly the Neutrality Thesis shifts the burden of ontological commitment *away* from the first-order quantifiers just as much as it does for the second order quantifiers. Indeed, why *shouldn't* the Neutrality Thesis apply just as well at the first order?

The case for a conception of first-order quantification that is not entangled with ontological commitment is forcefully made by Routley in "On What There Is Not" [29]. Routley undertakes the project of laying the groundwork for an ontology replete with all sorts of objects, only some of which are characterized by *existence* or *being*. But more empirically or nominalistically inclined philosophers can still appreciate the strength of his arguments against the identification of existence with quantification (or "quantifiability"), an identification which according to Routley is "as false as it is simple" [29, p. 169]. Renouncing the identification goes hand in hand with a conception of first-order quantification that is existentially neutral, as Wright also would have it. The challenge is to make formal sense of such a notion of neutrality. It is relatively easy to provide an account of existentially neutral quantification in variable-domain modal contexts (in which quantifiers are taken to range over the union of all possible domains, with existence, i.e., existence at the *actual* world, to be settled separately; see [23, pp. 341–42] for instance). It is not quite as clear how to make sense of this notion in non-modal contexts, a question that will take center stage in what follows.

Before we start developing our account of the semantic indeterminacy of first-order quantification, it should be mentioned that there is a robust tradition originating with [5] that also views second-order quantifiers as not more ontologically committed than their first-order counterparts. Second-order quantifiers like the one occurring in $\exists X\Phi(X)$ are analyzed as quantifying over *pluralities*, i.e., as asserting that "there are some x's such that Φ," where the plural locution "there are some x's such that ..." is in important cases demonstrably not paraphrasable at the first order. Insightful as this line of thought might be for philosophy of logic, it is not germane to the present concerns because it is based on highlighting special features of second-order quantifiers (*viz.*, that they refer to pluralities), features that they do not share with the first-order quantifiers, whereas the emphasis is here on the features that first-order quantifiers share with the second-order ones.

9.3 Quantifiers as Second-Order Predicates

A formula such as $\varphi(x)$, in which the variable x occurs free, can be viewed as expressing a *predicate* over some given first-order domain D_1 of objects, and in particular as expressing the predicate under which all and only the objects fall that satisfy the formula $\varphi(x)$ in D_1 (we follow the terminological conventions of some neo-Fregean approaches in referring to subsets of the domain as "predicates" and the formulas that correspond to them as "predicate expressions"). For instance, the formula $Px\&\neg Qx$ expresses the predicate under which all and only the objects fall that have the property denoted by P but lack the property denoted by Q. Similarly, formulas with more than one variable express relations over D_1, etc. Plausible as this view might appear, it has nonetheless noteworthy implications for the proper understanding of the quantifiers.

The essential function of a quantifier symbol such as $\exists x$ or $\forall x$ is to combine with a formula $\varphi(x)$ to obtain a sentence. It is accordingly natural to characterize first-order quantifiers—*viz.*, the operators denoted by $\exists x$ or $\forall x$—as second-order predicates, i.e., predicates of predicates. On this view, for instance, the existential quantifier applies to all and only the non-empty predicates, returning value *true* when, and only when, the predicate to which it is applied is non-empty; and dually the universal quantifier applies to a predicate if and only if that predicate is equi-extensional with the whole domain D_1. This is in fact a view that was already advocated by [8, §21], before being given rigorous mathematical formulation in the theory of generalized quantifiers initiated by [16, 20], and [19]. According to the theory, in fact, any permutation invariant second-order predicate (i.e., in purely extensional terms, any collection of subsets of D_1 containing a subset if and only if it contains any other subset equinumerous to it) gives rise to a quantifier. We have thus a vast collection of quantifiers beyond those denoted by $\exists$ and $\forall$, including for instance the quantifiers corresponding to "there are exactly k" (for $k \geq 0$), "there are infinitely many," or "for all but finitely many" (and many others).

Once first-order quantifiers are construed as second-order predicates, Henkin's development of non-standard interpretations for second-order logic can be transferred at the first-order level as well. On Henkin's non-standard interpretations second-order quantifiers are taken to range over collections of predicates that may fall short of the true power-set of the domain. For instance, they may be taken to range over the collection of predicates that can be defined using a given, limited, stock of resources, etc. Such non-standard interpretations are referred to as "general," in that they arise by relaxing, in a natural way, the requirement that $\forall X$ and $\exists X$ range over the collection of *all* predicates over the first-order domain D_1. On the general interpretation, second-order logic is co-interpretable with an appropriately designed multi-sorted first-order logic, and therefore its expressive power is vastly inferior to the standard case.

In spite of the fact that Henkin's groundbreaking approach has been available for a long time, the theory of generalized quantifiers still shares the bias towards standard interpretations. The theory characterizes a first-order quantifiers as a predicate over the full power-set of the first-order domain D_1 (see, e.g., Peters and Westerståhl [22]). But it is only natural to take this view one step further, and recognize that first-order quantifiers—just like their second-order counterparts—are open to general interpretations. Such interpretations would supply, beside a first-order domain D_1, also a second-order domain D_2 of subsets of D_1. The existential quantifier would then select, from among the members of D_2, those that are non-empty, and dually the universal quantifier would select the members of D_2 that are co-extensive with the whole domain D_1. The view neatly extends to other first-order quantifiers: the quantifier "there are infinitely many" would select those members whose cardinality is greater than or equal to the cardinality of the natural numbers, etc. Although non-standard interpretations for the first-order quantifiers have been around at least since [31] and [12], this particular extension of Henkin's general interpretations appears to have gone mostly unnoticed until Antonelli [1].

Somewhat more formally, we can specify a general semantics for a first-order language (e.g., a language such as that of classical first-order logic with $\exists$ and its dual $\forall$), as follows. A model $\mathfrak{M}$ provides a non-empty first-order domain D_1 along with a collection D_2 of *non-empty* subsets of D_1. Truth on such an interpretation can then be defined by saying, for instance, that $\exists x\varphi(x)$ is true in $\mathfrak{M}$ if and only if the extension $\{x \in D_1 : \varphi(x)\}$ of $\varphi(x)$ is a non-empty member of D_2 (a rigorous definition will first proceed to define satisfaction of an open formulas by an assignment to the variables; see Antonelli [3] for details). Notice that on this account if we were to define $\forall$ in a similar way, then duality with $\exists$ is lost. That is, if we define $\forall x\varphi(x)$ to be true in $\mathfrak{M}$ if and only if the extension of $\varphi(x)$ is a member of D_2 that is equi-extensional with D_1, then $\forall$ turns out no longer to be the dual quantifier of $\exists$. Duality can be retained by taking $\exists$ as primitive and abbreviating $\neg\exists x\neg\varphi$ by $\forall x\varphi(x)$: then $\forall x\varphi(x)$ is true in $\mathfrak{M}$ precisely when the extension $\{x \in D_1 : \neg\varphi(x)\}$ of $\neg\varphi(x)$ is a either not a member of D_2, or else such an extension is empty.

While the full technical details are explored in Antonelli [3], we have enough to point out some of the ramifications of this account for the philosophical use of quantifiers as carriers of ontological commitment. These ramifications will be explored more in detail in later sections of this paper, but for now we point out that, as Quine's thesis can be decomposed into separate necessity and sufficiency claims, the possibility of providing general interpretations for the first-order quantifiers has different implications for each half.

According to the necessity thesis, being the value of a bound variable is necessary for ontological commitment: objects to which we are ontologically committed are available as values for bound variables and therefore ground the truth of the corresponding existentially quantified statement. According to the necessity thesis, if we are ontologically committed to a, and a satisfies $\varphi(x)$, then $\exists x\varphi(x)$ must be true, because of the commitment to one of its instances. But notice (a point to which we will return in more detail) that on the general interpretation of first-order quantifiers, the necessity thesis fails, for there will be interpretations in which $\varphi(x)$ is satisfied by a, but the extension $\{x \in D_1 : \varphi(x)\}$, while non-empty, will not be among the subsets in the collection D_2 used in stating the truth conditions for existentially quantified statements.

While the failure of the necessity thesis on the general interpretation is clear-cut, things are somewhat murkier in the case of the converse. According to the sufficiency thesis, if the quantified statement $\exists x\varphi(x)$ is true, then there must be objects instantiating $\varphi(x)$. This conclusion seems difficult to avoid, even on the general interpretation, for if the extension $\{x \in D_1 : \varphi(x)\}$ is a member of D_2 to which the quantifier applies, then in particular such an extension must be non-empty, and therefore there must be objects satisfying $\varphi(x)$ and instantiating the bound variable. We postpone further discussion of this point until later, when we will see that there is at least a way to make the notion of a failure of the sufficiency thesis coherent.

9.4 The Case for "there is" and "for all"

We mentioned that Quine's linking of semantics and ontology in his characterization of ontological commitment presupposes a special role for the two quantifiers "there is" and "for all." This is true regardless of whether one approaches the characterization in order to support it or to undermine it. There is indeed a question as to why these two quantifiers, among many, have emerged to play such a central role, not only in philosophical disputes, but throughout the spectrum of human inquiry. And of course, any argument directed at explaining how that specific role bears upon ontological commitment runs the risk of being made irrelevant if other quantifiers are available to claim the same fundamental function. This section tells a story as to why these two quantifier are so prominent, a story that takes us on a detour through binary quantifiers and determiners (which, in this section, are only construed on their standard semantics).

We saw that from the point of view of the theory of generalized quantifiers, first-order quantifiers are to be identified with predicates over the power-set of the first-order domain D_1. But besides "unary" quantifiers applying to one predicate at a time, such as those expressed by $\exists$ and $\forall$, "binary" quantifiers applying to two predicates are also quite common, and in fact probably even more so, as they are ubiquitous in natural language, where they provide extensions for *determiners*. The Aristotelian quantifiers "All" and "Some" are in fact determiners, taking two predicates A and B as arguments (referred to as the "scope" and the "range" of the determiner, respectively) and returning propositions of the form "All A's are B's" or "Some A's are B's." But many more relations between predicates can be expressed by determiners, as revealed by even a cursory glance at the following partial list (from [22, p. 120]):

> Some, a, all, every, no, several, most, neither, the, John's, at least 10, all but 10, infinitely many, about 200, an even number of, between 5 and 10, most but not all, either fewer than 5 or more than 100, John's but not Mary's, at least one of most students', neither the red nor the green, ...

The fact that determiners play such an extensive role in ordinary language and communication is a sign of their fundamental nature. Determiners express the fact that the scope and the range are related in some particular way: this is just another way of saying that their denotations are binary quantifiers. The Aristotelian determiner "All" expresses that the scope and the range are related by inclusion, "Some" expresses that the scope and the range have non-empty intersection, etc.

It is natural to ask what a most generic, or *weakest*, determiner would look like: such a determiner would simply be expressing the fact that the scope and the range are related *some way or other*, i.e., expressing the existence of a relation having the scope as its domain and the range as co-domain. Let us define the *most general* determiner as denoting the binary quantifier holding between A and B precisely when there is some relation R relating objects in A to objects in B. Given some widely accepted set-theoretic assumptions, the existence of such a relation R is equivalent to the existence of a function f mapping objects in A to objects in B (clearly a function is a special kind of relation, and conversely any relation can be refined to a function by selecting for each $a \in A$ a unique $b \in B$ such that Rab). In order to simply matters, we will accordingly denote the most general determiner by Q^f, where the binary quantifier Q^f holds of A and B if and only if there exists $f : A \to B$.

It would indeed appear that the denotations of many, perhaps all, natural-language determiners can then be obtained by placing further restrictions on the function f mapping the scope into the range of the determiner. For instance the Aristotelian determiner "Some" can be characterized in this way by saying that "Some A's are B's" is true precisely when there is a function $f : A \to B$ having a fixed point, i.e., an object $a \in A$ such that $f(a) = a$. Similarly, "All" can be characterized by saying that there is a function $f : A \to B$ such that $f(a) = a$ for each $a \in A$. The conjecture that many, perhaps all, natural language determiners can be obtained in this way (that is, by placing appropriate conditions on the function mapping the scope into the range) lends plausibility to the characterization of Q^f as a most general, or weakest, determiner, for then other determiners can be obtained by refining and strengthening it.

It would seem natural to conjecture that the expressive power of a quantifier varies in accordance with the strength of the restrictions imposed on the function f relating the scope and the range of the corresponding determiner. This amounts to saying that if quantifiers Q_1 and Q_2 are defined by imposing given restrictions Φ_1 and Φ_2 on f:

$$Q_1(A,B) \Leftrightarrow \exists f\big[f : A \to B \ \& \ \Phi_1(f)\big],$$
$$Q_2(A,B) \Leftrightarrow \exists f\big[f : A \to B \ \& \ \Phi_2(f)\big],$$

and $\Phi_1(f)$ implies $\Phi_2(f)$ for any f, then the expressive power of Q_1 is at least as great as that of Q_2. We will not specify precisely how expressive power is to be measured, but a necessary condition for the expressive power of Q_1 to be at least as great as that of Q_2 would seem to be that Q_1 interprets Q_2 (relative to some given background language). But this conjecture fails, along with its converse, showing that the expressive power of the quantifier is independent of the strength on the condition Φ_1 or Φ_2. Consider the following three binary quantifiers obtained by imposing increasingly stronger restrictions on the function f:

$$Q^f(A, B) \Leftrightarrow \exists f : A \to B,$$

$$Q^1(A, B) \Leftrightarrow \exists f : A \to B \;\; \& \;\; f \text{ injective},$$

$$Q^=(A, B) \Leftrightarrow \exists f : A \to B \;\; \& \;\; \forall a \in A : f(a) = a.$$

Clearly $Q^=(A, B)$ implies $Q^1(A, B)$, since identity is injective, and $Q^1(A, B)$ implies $Q^f(A, B)$. But $Q^1(A, B)$ is strictly stronger than the other two, and by far. To see this, observe that the *weakest* of the three quantifiers above, *viz.*, Q^f, is equivalent, over a weak logic comprising an identically empty predicate $\varnothing$ (i.e., a name for the empty set, such as $x \neq x$) and Boolean operators, to each of the ordinary quantifiers $\exists$ and $\forall$. For clearly, there exists a function $f : A \to B$ (which is what Q^f expresses) precisely when $B = \varnothing$ implies $A = \varnothing$, and obviously "if B is empty then so is A" is expressible using $\exists$ or $\forall$. Conversely, $\exists x Ax$ can be expressed as $\neg Q^f(A, \varnothing)$ and dually $\forall x Ax$ as $Q^f(\neg A, \varnothing)$. Thus the logic with Q^f as its only quantifier is essentially the same as ordinary first-order logic. Moreover, $Q^=$ is just the Aristotelian determiner "All A's are B's," or $A \subseteq B$, which is also expressively equivalent to first-order logic, as one easily sees. Thus the logic with $Q^=$ is also equivalent to first-order logic. However, the intermediate quantifier Q^1 is much more expressive than standard first-order logic. In fact Q^1 interprets well-known cardinality quantifiers such as Rescher's or Härtig's, and so it can be used, e.g., to provide a categorical axiomatization of arithmetic (a result that goes back to Yasuhara [35]; see also Antonelli [2]).

This detour on binary quantifiers and determiners gave us a characterization of the fundamental role played by the unary quantifiers $\exists$ and $\forall$. Our story identifies the general form of the determiner as expressing the existence of a functional relation between its two arguments, the scope and the range. When no further constraints are imposed on such a functional relation, we have a basic, most general, and arguably most natural form of the determiner. But as shown above such a form, in turn, is expressively equivalent to the two quantifiers, "there is" and "for all," thereby giving us at least the beginning of an insight into their fundamental nature and crucial role.

9.5 Varieties of General Interpretations

In this section we return to non-standard interpretations for first-order quantifiers, and specifically for those denoted by $\exists$ and $\forall$. As we will see, such interpretations come in many different kinds. For our present purposes by "first-order logic" we mean the first-order language obtained from a given stock of extra-logical symbols by means of truth-functional connectives and the quantifier $\exists$ (with $\forall$ defined as the dual of $\exists$). The contrast here is with first-order languages including arbitrary first-order quantifiers besides (or perhaps even instead of) these two, a general case that is given fuller treatment in Antonelli [3]. We have seen that on the general semantics for first-order logic, a model $\mathfrak{M}$ supplies a non-empty domain D_1 along with a collection D_2 of subsets of D_1. The existential quantifier is interpreted relative to the second-order domain D_2: $\exists x\varphi(x)$ is true in $\mathfrak{M}$ if and only if the extension $\{x \in D_1 : \varphi(x)\}$ is a non-empty member of D_2, and since we take $\forall$ as the dual quantifier, the definition at the same time fixes the truth conditions for universally quantified sentences as we saw in Sect. 9.3.

Consideration of general interpretations can help shed light on at least half of the Quinean thesis, that instantiating a bound variable—i.e., falling under a predicate which in turn falls under the quantifier—is necessary for ontological commitment. General interpretations make it clear that in this respect necessity fails: we can be committed to entities having certain properties without such entities being available as values of—truth-makers for—the corresponding existentially quantified sentences. In order to bring general interpretations to bear also on the other half of the Quinean thesis, sufficiency, we need to consider a class of interpretations that appear at first sight to be of quite a different kind.

There is a deep connection between non-standard models satisfying certain further constraints (to be specified in a moment) and the kind of inner-outer domain interpretations originally proposed for positive free logic (see [14]). Positive free logic was developed by [13] to include non-referring terms in such a way that not all atomic predications involving such terms are automatically false. Inner-outer domain interpretations accomplish this by providing *two* domains of objects: a possibly empty *inner* domain I and an *outer* domain O, with $I \subseteq O$. The inner domain is intended to comprise "existing" objects, those towards which we carry ontological commitment, whereas the outer domain provides denotations for non-referring terms. The extension of an atomic predicate might comprise objects from either domain, thereby allowing some predications involving non-referring terms to be true. Of course, the quantifiers $\exists$ and $\forall$ are then taken to range over the inner domain.

The connection between outer-domain models and general interpretations can be stated as follows: for every inner-outer domain model $\mathfrak{M}$ there exists a generalized model $\mathfrak{N}$ verifying the same sentences. In fact, for a given inner-outer domain model $\mathfrak{M} = (I, O)$ the corresponding generalized model can be obtained by putting $D_1 = O$ and defining D_2 to consist of all those subsets of D_1 that contain at least one member of the inner domain I. It follows that if a sentence φ is true in all general models then it is true in all outer-domain model, and hence the logic of general models is contained in positive free logic. The converse is not true: not all general models are equivalent to inner-outer domain models; however, it is possible to identify a condition on which a general model is, in fact, equivalent to an inner-outer domain model. The condition simply requires that if D_2 contains a predicate P having non-empty overlap with a predicate Q, then Q is also a member of D_2 (details are given in Antonelli [3]). We will return to this fact in the next section.

It is interesting perhaps that some non-standard interpretations of the two quantifiers "there is" and "for all" recover, in fact, the standard reading (the one on which they are taken to range over the full power set of D_1), without having the quantifiers apply to the full power-set of the first-order domain. Consider for instance a standard interpretation $\mathfrak{M}$ for a first-order language, with domain D_1, and let D_2 consists of all non-empty subsets of D_1 that are *first-order definable* in $\mathfrak{M}$ by a formula φ (possibly involving parameters). We can then form the "hull" of $\mathfrak{M}$, i.e., the non-standard interpretation $\mathfrak{N}$ having D_1 and D_2 as its first- and second-order domains, respectively (i.e., the hull of $\mathfrak{M}$ has the same first-order domain, but a "lean" second-order domain consisting only of the subsets that are definable in $\mathfrak{M}$). It is not obvious that the quantifiers have the same meaning in the original model as they do in the leaner hull. It is therefore somewhat unexpected that exactly the same formulas are satisfied by the same members of D_1 in the two interpretations (see Antonelli [3]). This could be construed as evidence that the standard interpretation "overshoots" the target by requiring that the quantifiers range over the full power set, when in fact much less than that is sufficient to recover the ordinary construal.

The fact that any standard first-order model is equivalent to its hull shows that there is some kind of "reflective equilibrium" that seems to characterize genuinely first-order notions *vis-à-vis* second-order ones. For we have seen that the difference between first- and second-order notions is *not* in the fact that, as one might be tempted to assume, the second-order notions, but not first-order ones, are semantically sensitive to interpretations whose second-order domains falls short of the full power set. Rather, a potentially useful criterion for demarcating first-order and second-order notions is whether these notions can be characterized as invariant upon transition to hulls, i.e., upon restricting the interpretations to those having a second-order domain comprised only of definable subsets. Here of course "definable" means "definable using those very notions:" this apparent circularity points in fact to a the reflective equilibrium characterizing first-order notions.

Now, of course, one could worry about the circularity, and question the significance to be attached to the results just mentioned, because after all, in order to be able to specify the class D_2 of all first-order definable subsets of D_1 (i.e., those definable on the standard interpretation), the general interpretation obtained depends on a *prior* understanding of the standard meaning of the quantifiers. But there are ways around this, although they require some fair amount of detail in building up the collection of definable subsets "from below," as it were, rather than by the roundabout device of identifying such subsets wholesale in terms of the standard satisfaction relation. The details of the procedure are given in the companion paper Antonelli [3] already mentioned, but the main idea is to identify the appropriate collection of Gödel operations (first introduced in [9] in connection with the constructible universe) and then characterize the collection of all definable subsets of the domain as the closure of the class of primitive relations over D_1 (including identity) under those functions. The definable subsets are thus identified *directly* without resorting to the standard notion of the quantifier.

9.6 The Quinean Thesis: Necessity and Sufficiency

We now have the tools necessary to re-assess the Quinean thesis that "to be is to be the value of a bound variable," as composed by the two distinct claims: that being the value of a bound variable is *sufficient* for ontological commitment, and the converse claim that being the value of a bound variable is *necessary* for ontological commitment.

The possibility of non-standard interpretations for first-order quantifiers, and particularly for $\exists$ and $\forall$, makes it clear that the necessity thesis fails in a crucial sense. This is because on the general interpretation of the first-order quantifiers, $\exists$ and $\forall$ apply only over a subset of the true power set of D_1. There might then be members of D_1 that are not captured by the two quantifiers: they, so to speak, lie beyond the quantifiers' reach. This can happen, as mentioned, if a formula $\varphi(x)$ is satisfied by some object in D_1 but the extension $\{x \in D_1 : \varphi(x)\}$ is not a member of the second-order domain D_2, thus making $\exists x\varphi(x)$ false on the interpretation: in such a case the necessity thesis fails. Thus, on the general interpretation of the quantifiers, being the value of a variable is not necessary for ontological commitment, in that some members of D_1 are not available as possible instantiation of the variable in $\exists x\varphi(x)$ and $\forall x\varphi(x)$. As an immediate consequence of this fact, we see that inference patterns that take the form of existential generalization fail on this semantics. That is because instances of the form $\varphi(a)$ might be true, but their existential generalizations, $\exists x\varphi(x)$, might not, as we just observed, so that existential generalization is not truth-preserving.

We also saw that there is a deep connection between the general interpretation of the quantifiers "there is" and "for all" and the inner-outer domain semantics: every inner-outer domain corresponds to an equivalent general model, and conversely every general model satisfying the condition given in Sect. 9.5 is equivalent to an inner-outer domain model. But this technical fact should *not* be taken to mean that the logic of the generalized "there is" and "for all" is in fact just a variant of free logic. In fact the two are approaches have very different inspirations and outcomes. Free logic was originally developed as the logic of *non-referring terms*, be they atomic terms such as "Pegasus" or "Bellerophon," or complex terms such as the empty descriptions "the winged horse" or "the golden mountain." On the contrary, the general interpretation of the quantifiers has nothing to do with "non-existent" objects (of which, we submit, there are none), but rather with the possibility that some particular objects (on some interpretations) might lie beyond the reach of the first-order quantifiers. This would appear to be a more natural construal of the failure of existential generalization than the one given by free logic, at least in its positive version, requiring formulas to be satisfied by objects that are, by the lights of the logic itself, "non-existent."

A parallel analysis can be given in the second-order case. We mentioned that second-order quantifiers can be given a general interpretation on which they are taken to denote some subset of the double power-set of D_1, as first pointed out by Henkin. And in fact, any second-order model in which some instance of second-order comprehension *fails* must be of this kind, for then there will be complex predicates expressions $\Phi(x)$ that fail to correspond to some subset X of D_1 falling within the range of the second-order quantifiers $\exists X$ or $\forall X$. This is possible because although of course the extension of $\Phi(x)$ will be among the members of the true power-set of D_1 it will not, in general, be among those over which the second-order quantifiers range (systems with restricted second-order comprehension play an important role in the foundations of mathematics). It is of course possible to regard such complex predicates—in analogy to the first-order case—as *non-referring*: on this account second-order logic, on the general interpretation, could be viewed as the free logic of these non-referring predicates. But it is more natural, and customary, to regard the general interpretation as allowing predicates that exceed the grasp of the second-order quantifiers. The parallel construal of first-order quantifiers is similarly more natural than the extravagant ontology required by positive free logic.

We now turn to the other half of the Quinean thesis, sufficiency. Are there any more insights to be gained from considering the sufficiency thesis in the light of this understanding of the quantifiers, comparable to those delivered by consideration of the necessity thesis? As we have seen in Sect. 9.2, several arguments have traditionally been put forward to defuse the sufficiency thesis, but only as it applies to *second-order* quantifiers. Both Prior's Insight and Wright's Neutrality Thesis are squarely aimed at the second-order case, when as it should now be evident there is nothing specifically second-order about either one of them, and they equally well could—and *should*—be applied at the first order. This is especially clear in the case of what we called Prior's Insight: if ontological commitment is preserved through converse logical consequence, then the commitment of $\exists x\varphi(x)$ cannot exceed that of $\varphi(a)$, from which it follows, and since the latter is not ontologically committed, the sufficiency thesis fails.

A challenge remains, though. Given the generalized semantics of the quantifiers, is there any sense to be made of the failure of the sufficiency thesis? It would seem that the sufficiency thesis can only really fail in the presence of extravagant ontologies—be they Routley's rich ontology of non-existents or the more mathematically precise ontology of variable-domain quantified modal logic. But if we want to avoid the former while remaining in a non-modal context we need some account of how exactly the sufficiency thesis is supposed to fail: being the value of a bound variable, while clearly not necessary, would seem to be at least sufficient for *actual* existence. In other words, even if the non-standard semantics makes it clear that objects in D_1 might outstrip the grasp of $\exists$ and $\forall$, it still *does* follow that the truth of $\exists x\varphi(x)$ carries with it commitment to the extension of $\varphi(x)$'s being non-empty. For $\exists x\varphi(x)$ is true on an interpretation precisely when the extension of $\varphi(x)$ is among the predicates in D_2 and such that the quantifier applies to it, and this last condition requires that some object from D_1 must be a member of the extension of $\varphi(x)$.

In order to address this point, let us first observe that the notion of a failure of the sufficiency thesis is at least *coherent*, in that it can be realized in a model. In fact, such a model can be obtained by combining the general interpretation of the first-order quantifiers with a device derived from the realization that existence is no longer the exclusive purview of the quantifier, i.e., that existence is an extra-logical and extra-semantical notion (as also [4] would have it), and that therefore it needs to be expressed by extra-logical means. We can implement this idea by recourse to a form of outer-domain semantics, as follows. Given a first-order domain D_1, identify a subset I of D_1 as the inner domain of individuals to whose existence we are ontologically committed, and let the existential quantifier $\exists$ range over some collection of non-empty subsets of D_1 (note: *not* necessarily subsets of I), and dually for $\forall$. This means that a model is now identified with a triple (D_1, D_2, I), where $I \subseteq D_1$ and $D_2 \subseteq \mathcal{P}(D_1)$. In such a model, then, being the value of a bound variable is neither necessary nor sufficient for ontological commitment. It is not necessary because, as before, some member of D_1 might instantiate a formula whose extension fails to be in D_2, so that the existential closure of that formula fails to be true. But the sufficiency thesis fails as well, in that some formula $\varphi(x)$ might have a non-empty extension in D_2 comprised only of members of D_1 that fall outside of I. The corresponding existential $\exists x\varphi(x)$ would then be true, but its instances would not meet the extra-logical standard needed for ontological commitment. Therefore, to the extent to which the inner domain I represents the ontological commitments of the interpretation, being the value of a bound variable does not guarantee being the object of such a commitment.

But this argument against the sufficiency thesis is importantly different from the one we gave against the necessity thesis. The latter is a "for all we know" argument, in the sense that, *for all we know*, there might really be objects exceeding the grasp of our ordinary first-order quantifiers. On *some* interpretations these "outlying objects" will then be logically inert, thereby making the particular interpretation of the quantifiers completely transparent to us. In particular, this will be the case for those interpretations in which quantification supplies the only route providing access to the objects of the domain. But there will also be interpretations in which these objects will *not* be logically inert: one lesson we can learn from the proponents of free logic, is that singular terms—atomic or complex—can also provide an access route to objects that do not fall within the reach of the quantifiers. We need not characterize these objects as "non-existent," but we need to recognize that ontological commitment to such objects is prior to, and independent of, the semantics of the quantifiers. This is also the reason why non-standard interpretations cannot be ruled out, at least not on purely linguistic grounds: they might be completely transparent to users of the language, who would then need to resort to other extra-semantical means to assess questions of existence.

On the other hand, the argument establishing the coherence of the notion of a failure of the sufficiency thesis only provides a model. The model is not quite viable as a model of anything resembling actuality, unless one subscribes to extravagant ontologies (which we are trying to avoid). Non-existent objects should be just that—non-existent—and therefore no model that countenances such objects or their *simulacra* can lay any claim to realistic plausibility. But the model in which sufficiency fails still plays a role, perhaps a crucial one. For the model makes clear what the direction of the explanation ought to be: it's not the quantifier that *enacts*, so to speak, the ontological commitment, but rather it's the semantics for the quantifier that depends in the first instance on a *prior* selection of a domain D_1 of quantification, *as well as* an accompanying a second-order domain D_2, just as Wright's Neutrality thesis would require, once properly extended to include first-order quantifiers.

9.7 Conclusion

Being the value of a bound variable is neither necessary nor (in an importantly different sense) sufficient for ontological commitment. Consideration of non-standard models for the existential and universal quantifiers reveals that the possibility that there might be objects that lie beyond the reach of such quantifiers cannot be ruled out, or at least not on purely semantical grounds. Whether the language so interpreted has the expressive means to discriminate such cases depends on the details of the interpretation (a rich enough interpretation might be indistinguishable from the standard one—as the one comprised of definable subsets shows). Thus the necessity thesis indeed easily fails.

But the sufficiency thesis is also far from unassailable, in the sense that its failure is at least a coherent notion. This is apparent by turning on its head the free logician's expedient of an outer domain: instead of constraining the quantifiers to range over the inner domain, as free logicians do, we allow the truth of some existentially quantified sentences whose only witnesses lie outside the inner domain, but retain membership in the inner domain as a measure of ontological commitment. Thus the ontological commitments that accompany our use of the first-order quantifiers, such as they are, are in fact *dependent* upon our prior selection not just of a first-order domain, D_1, but also of a second-order domain D_2.

Of course, this conclusion still leaves open the question of exactly *how* those ontological commitments are to be established in the first place, i.e., how we go about the relevant domain selections. We are thus brought back to a broader Quinean conception according to which ontological commitments are in some way bundled up in a more or less holistic manner with our linguistic practices (a conception that Quine himself would have found appealing, as we noticed at the beginning). It's just the atomistic attempt to individuate one particular linguistic component—the quantifiers—as the specific *locus* of the commitments that fails. The result is that we are left facing a radical indeterminacy in the semantics of the quantifier, not dissimilar to the indeterminacy of numerical notions brought about by non-standard model for arithmetic. In both cases determination of whether language suffices to pin down the intended model requires access to an external, independent viewpoint unavailable within the limited expressive resources of first-order languages.

From a formal point of view, the upshot of the discussion is that first-order quantifiers are just as semantically sensitive to general interpretations with a non-standard second-order domain as their second-order counterparts. The realization cuts both ways, though. On the one hand it makes clear that indeterminacy in the semantics of the first-order quantifiers cannot be addressed simply by fixing a first-order domain D_1: a second-order domain D_2 needs to be specified as well, just as for second-order logic. But on the other hand, any reservations one might have concerning the ontological commitments of second-order logic can be assuaged by the fact that those are the same as in the first-order case, which has long been considered ontologically innocent; and this last realization can contribute to the establishment of second-order logic on the same safe footing as first-order logic.

Acknowledgements I am grateful to Sean Ebels-Duggan, Robbie Hirsch, Elaine Landry, and Adam Sennet for helpful comments.

References

1. Antonelli, G.A. 2007. Free quantification and logical invariance. *Rivista di estetica* (special issue "Il significato eluso. Saggi in onore di Diego Marconi," ed. M. Andronico, A. Paternoster, and A. Voltolini), 33(1): 61–73.
2. Antonelli, G. A. 2010. Numerical abstraction via the Frege quantifier. *Notre Dame Journal of Formal Logic* 51(2): 161–79.
3. Antonelli, G.A. 2013. On the general interpretation of first-order quantifiers. *Review of Symbolic Logic* 6: 637–658.
4. Azzouni, J. 2004. *Deflating existential consequence: A case for nominalism*. Oxford/New York: Oxford University Press.
5. Boolos, G. 1984. To be is to be a value of a variable (or to be some values of some variables). *Journal of Philosophy* 81(8): 430–449.
6. Dutilh Novaes, C. 2011. Medieval theories of quantification. In *Encyclopedia of medieval philosophy*, ed. H. Lagerlund. Dordrecht/New York: Springer.
7. Fara, D., and G. Russell. 2011. *The Routledge companion to the philosophy of language*. New York: Routledge
8. Frege, G. 1893–1903. Grundgesetze der Arithmetik. *The basic laws of arithmetic: Exposition of the system*, ed. Jena Hermann Phole, Partial English Trans. Montgomery Furth. Berkeley: University of California Press, 1964
9. Gödel, K. 1940. *The consistency of the axiom of choice and the generalized continuum hypothesis with the axioms of set theory*. Annals of mathematics studies, vol. 3. Oxford/New York: Oxford University Press.
10. Henkin, L. 1950. Completeness in the theory of types. *Journal of Symbolic Logic* 15: 81–91.
11. Hirsch, E. 2002. Quantifier variance and realism. *Philosophical Issues* 12(1): 51–73.
12. Keisler, H.J. 1970. Logic with quantifier "there exists uncountably many". *Annals of Mathematical Logic* 1: 1–93
13. Lambert, K. 1963. Existential import revisited. *Notre Dame Journal of Formal Logic* 4(4): 288–292.
14. Leblanc, H., and R.H. Thomason. 1968. Completeness theorems for some presupposition-free logics. *Fundamenta Mathematicæ* 62: 125–54.
15. Leng, M., A. Paseau, and M. Potter (eds.). 2007. *Mathematical knowledge*. Oxford/New York: Oxford University Press.
16. Lindström, P. 1966. First-order predicate logic with generalized quantifiers. *Theoria* 32: 186–95
17. Linnebo, Ø. 2011. Second-order logic. In *The continuum companion to philosophical logic*, Ch. 4, ed. L. Horsten and R. Pettigrew. London/New York: Continuum.
18. Linnebo, Ø. 2012. Plural quantification. In *The stanford encyclopedia of philosophy*, ed. Edward N. Zalta. http://plato.stanford.edu/archives/sum2012/entries/plural-quant/.
19. Montague, R. 1974. English as a formal language. In *Formal philosophy*, ed. R.H. Thomason. Oxford/New York: Oxford University Press.
20. Mostowski, A. 1957. On a generalization of quantifiers. *Fundamenta Mathematicæ* 44: 12–36.
21. Parsons, T. 1970. Various extensional notions of ontological commitment. *Philosophical Studies* 21(5): 65–74.
22. Peters, S., and D. Westerståhl. 2006. *Quantifiers in language and logic*. Oxford/New York: Oxford University Press.
23. Priest, G. 2008. *An introduction to non-classical logics. From if to is*. Cambridge: Cambridge University Press.

24. Prior, A.N. 1971. *Objects of thought*, ed. P.T. Geach and A.J.P. Kenny, ix–175. Oxford: Clarendon Press.
25. Quine, W.V.O. 1948. On what there is. *Review of Metaphysics* 2(5): 21–38.
26. Quine, W.V.O. 1970. *Philosophy of logic*, 2nd ed. Englewood Cliffs: Prentice-Hall. 1986, by Harvard University Press
27. Rayo, A. 2007. Ontological commitment. *Philosophy Compass* 2/3: 428–444.
28. Rayo, A, and Yablo, S. 2001. Nominalism through de-nominalization. *Nôus* 35(1): 74–92.
29. Routley, R. 1982. On what there is not. *Philosophy and Phenomenological Research* 43(2): 151–177.
30. Sider, T. 2007. Neo-fregeanism and quantifier variance. *Aristotelian Society Supplementary Volume* 8(1): 201–232.
31. Thomason, R.H., and D.R. Johnson. 1969. Predicate calculus with free quantifier variables. *The Journal of Symbolic Logic* 34(1): 1–7
32. Thomasson, A.L. 2008. Existence questions. *Philosophical Studies* 141(1): 63–78.
33. Van Inwagen, P. 1987. When are objects parts? *Philosophical Perspectives* vol. *I: Metaphysics* 1: 21–47.
34. Wright, C. 2007. *On quantifying into predicate position*. ed. M. Leng, A. Paseau, and M. Potter [15], 150–174.
35. Yasuhara, M. 1969. Incompleteness of L_p languages. *Fundamenta Mathematicae* 66: 147–152.

Chapter 10
Chalmers, Quantifier Variance and Mathematicians' Freedom

Sharon Berry

Abstract Philosophers of mathematics have been much struck by mathematicians' apparent freedom to introduce new kinds of mathematical objects, such as complex numbers, sets and the objects and arrows of category theory. In this paper, I explore a way of using recent work on quantifier variance to explain this apparent freedom to introduce theories about new kinds of mathematical objects. In Ontological Antirealism, David Chalmers sketches a method for describing a class of alternative quantifier senses which are more ontologically profligate than our own using appeals to set theoretic models. I suggest a modification of this method which frees it of certain arbitrary limitations on size, by replacing appeals to set theory with appeals to an (independently motivated) notion of broadly logical possibility. Once amended in this way, Chalmers' technique allows us to flesh out a Neo-Carnapian explanation for mathematicians' freedom to introduce new kinds of mathematical objects which avoids some major problems for existing accounts.

10.1 Introduction

Philosophers of mathematics have been much struck by mathematicians' apparent freedom to introduce new kinds of mathematical objects, such as complex numbers, sets, and the objects and arrows of category theory. For example, in a recent *Australasian Journal of Philosophy* paper Julian Cole writes, "Reflecting on my experiences as a research mathematician, three things stand out. First, the frequency and intellectual ease with which I endorsed existential pure mathematical statements and referred to mathematical entities. Second, the freedom I felt I had to introduce a new mathematical theory whose variables ranged over any mathematical entities

S. Berry (✉)
School of Philosophy, Australian National University, Coombs Building 9, Fellows Road, Canberra, ACT 0200, Australia
e-mail: seberry@invariant.org

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_10

I wished, provided it served a legitimate mathematical purpose. And third, the authority I felt I had to engage in both types of activities. Most mathematicians will recognize these features of their everyday mathematical lives."[5].

Various major strategies for explaining this apparent freedom to introduce new mathematical objects have been discussed in the literature, and each faces significant known problems. Recent work on quantifier variance suggests a way of explaining mathematicians' freedom which avoids these problems. On this neo-Carnapian approach, when mathematicians adopt suitably coherent hypotheses characterising new types of objects (as above), this can be regarded as an act of *ontologically inflationary stipulative definition* which gets us to start talking in terms of new objects. Such stipulations give meaning to newly coined mathematical expressions. They can also change the meaning[1] of the existential and universal quantifiers[2] as required to make the relevant stipulations express truths.

However, significant doubts have been raised about the intelligibility of the alternative, more ontologically profligate, quantifier meanings required by this story. Everyone allows that there can be quantifier restrictions, as when someone says, "All the beers are in the fridge." But, it is much more controversial to suggest that changes in language can allow "∃" to retain its usual inferential role, while taking on meanings which are not mere restrictions of some fundamental unrestricted notion of existence.

In this paper, I will defend the intelligibility of these alternative ontologically profligate quantifier meanings, by elucidating their contribution to truth conditions for whole sentences. In [4], David Chalmers suggests a way of describing a class of alternative quantifier meanings which are not mere quantifier restrictions by appealing to (something like) set theoretic models. I will suggest an alternate approach which avoids arbitrary limitations created by Chalmers' reliance on set theory. This approach replaces appeals to set theory with *fundamentally modal* claims about logical possibility. It allows us to clearly describe the behavior of the kind of alternative ontologically profligate meanings for "∃" which the Neo-Carnapian explanation of mathematicians' freedom requires. I will also sketch a story about how acts of mathematical stipulation can shift us between these various quantifier meanings in a suitable way.

[1] I will follow convention in describing changes in truth conditions for sentences which intuitively might be said to recognize new objects as changes in the meaning of the quantifier. However, I do not mean to commit myself to the view that in these situations it is the meaning of the quantifier which changes, rather than something like the Kaplanian content[3], or to the view that there are objects called 'meanings'.

[2] As we will see these changes in meaning will preserve the intuitive understanding of "there is" and "∃."

10.2 Alternative Approaches

Other theories have been offered to explain mathematicians' apparent freedom to introduce new kinds of objects. However, most of these approaches have had trouble satisfying at least one of the following, intuitively appealing, desiderata:

1. Avoid ruling out intuitively acceptable mathematical practices or positing arbitrary-looking limits on the mathematical universe.
2. Account for the apparent metaphysical necessity of mathematical truths.
3. Treat mathematical and non-mathematical existence claims in a similar fashion.

10.2.1 Limitative Approaches

Limitative approaches like classic set theoretic foundationalism interpret mathematicians as talking about objects drawn from a fixed, but large, mathematical universe. They explain mathematicians' freedom to introduce new kinds of mathematical objects by saying that all *acceptable* characterizations of mathematical structures can be understood as truly describing portions of this large universe. Thus, for example, set theoretic foundationalist versions of the limitative approach take acceptable characterisations of mathematical structures to be those which can be interpreted as truly describing some portion of the hierarchy of sets, and claim that when people talk about mathematical objects like numbers, graphs and ordered pairs they are really talking about suitable sets.

Limitative approaches face a problem with the first desideratum, in that they seem committed to either positing unprincipled-looking restrictions on what kinds of objects mathematicians can introduce or unprincipled-looking boundaries to the mathematical universe as a whole. To avoid arbitrariness, proponents of the limitative approach would like to say that all coherent *descriptions* of putative structures which mathematicians might choose to adopt[3] will be satisfied somewhere in the (fixed) mathematical universe.

The limitavist cannot justify this claim by saying that all *logically possible* structures are realized in this fixed mathematical universe. For, because of Russell-style paradoxes, it is intuitively logically possible to extend any supposedly complete collection of mathematical objects, in such a way as to generate a new structure which cannot be realized by any portion of the original mathematical universe.[4]

[3]Or at least all descriptions which are suitably 'internal', in the sense that they describe the internal structure of the relevant collection of mathematical objects without imposing constraints on the size of the universe as a whole.

[4]Intuitively, we can make sense of the notion of *all possible ways of choosing objects out of a collection*. Consider extending one's original mathematical universe by adding a layer of objects which 'witnesses' all possible ways of choosing from the original objects (in essence, adding a layer of classes to the original universe). For familiar Cantorian reasons, it would be logically

Proponents of limitative approaches might respond by saying that the mathematical universe can (nonetheless) contain witnesses to all coherent patterns of relationships *which human beings can describe in suitably mathematical vocabulary*. However, taking this line appears to involve positing an arbitrary-looking joint in reality. For why does the mathematical universe have this particular structure – despite the fact that (as discussed in the previous paragraph) it would be perfectly logically coherent for it to be somewhat larger, and (by hypothesis) it can't be the case that we accept some mathematical description of the intended structure of this total universe which explains why it is this large and no larger?[5,6]

10.2.2 Institutional/Social Constructive Approaches

Institutional/Social Constructive approaches like Cole's[5] take mathematical objects to be institutional facets of reality which "exist in virtue of collective agreement" and are, in some sense, created and "sustained in existence by a relevant group of people collectively recognizing or accepting their existence."[5]. Crucially, this kind of creation is not a matter of changing the meaning (or contextually determined content) of certain expressions. Rather, just as priests can bring marriages into being via ritual, so too mathematicians can ensure the existence of some suitable collection of mathematical objects just by choosing to accept certain existential claims about such objects. These views face problems with the second desideratum: it is prima facie difficult to square this account with the idea that mathematical statements can be timelessly and necessarily true.[7] Taking mathematical objects' existence to be grounded in social facts, in the same way that the existence of money or countries is grounded in social facts, also strikes many people as directly odd or counter-intuitive.

The social constructive approach also has serious difficulty handling the possibility that a pair of mathematical communities might accept incompatible mathemat-

impossible for any 1-1 relation R to pair up all the objects in this extended structure with objects in some portion of the original mathematical universe. Thus, it would seem that there is a logically possible structure which is not isomorphic to any portion of our supposed total mathematical universe.

[5]If we could describe the whole universe via suitably mathematical vocabulary, then we could also describe the larger structure which extends it by adding a layer of classes. Thus, there would be a coherent, suitably mathematical, description of a structure which is not isomorphic to any portion of the total mathematical universe.

[6]See Wright and Shapiro in [18] for more discussion of this worry.

[7]Cole deals with this worry by noting that standard acts of social construction (such as founding a company or granting an individual some important social status) can take effect retroactively. For example, he notes that sports authorities can retroactively rule that a player has been on the 'injured list' for the past 2 days, and he suggests that mathematical authorities can similarly rule that sets and numbers exist timelessly and amodally.

ical principles. As Boolos[2] emphasized, not all logically consistent hypotheses about abstracta are logically compatible with one another. Indeed, even when we restrict our attention to intuitively attractive theories which do not impose any kind of upper bound on the size of the universe, it remains possible to find consistent but incompatible pairs of theories, as Uziquiano recently demonstrated by showing the incompatibility of certain natural axioms for set theory with ur-elements and natural axioms for mereology.[8] Now what happens if some such pair of incompatible theories gets adopted within separate mathematical communities? For example, what happens if Chinese mathematicians introduce sets but American mathematicians simultaneously introduce parities? It would seem that (contra the Institutional/Social Constructivist account) both mathematical practices cannot simultaneously succeed in grounding the existence of mathematical objects sufficient to ensure the truth of their foundational principles.

10.2.3 Hypotheticalist and Fictionalist Approaches

Hypotheticalist approaches hold that the true logical form of a mathematical utterance 'ϕ' is something like, 'if *D* then ϕ', e.g., 'it is logically necessary that if D then ϕ' or 'if it were the case that D, then ϕ'[9] where *D* combines all relevant stipulations/assumptions/definitions about the mathematical objects accepted by the speaker. Such approaches bear some formal similarity to the proposal I will ultimately advocate, and have many attractive qualities. However, they do not meet the third desideratum, because the hypotheticalist takes mathematical existence claims to have a very different logical structure from existence claims about ordinary and scientific objects. In particular, sentences like, 'There is a city' and 'There is a number' share the same surface grammatical structure and play the same role in logical deductions, but would be taken to have a very different logical form

[8]For example, although it seems coherent to say that every object belongs to a set, and it seems coherent to say that every plurality of objects (including pluarlities of abstract objects like sets) has a mereological fusion, Uziquiano shows that one cannot *conjoin* a certain popular axiomatization of set theory with popular axioms of mereology on pain of logical contradiction[20]. Therefore, not all coherent stipulations can be consistently combined. In essence, the problem is that Uziquiano's principles of applied set theory and mereology both include claims about how their respective objects relate to *all other objects*, which imply incompatible consequences about the size of the universe as a whole. For example, Uziquiano notes a conflict between, "Atomicity: There are no objects whose parts all have further proper parts.", "Limitation of Size: Some objects form a set if and only if there is no 1-1 map from the entire universe into them." and some more commonplace axioms of set theory and mereology. Combining Atomicitiy with standard principles of mereology turns out to require that the universe as a whole to have size 2α for some cardinal α. But combining Limitation with standard axioms of set theory (which imply that *the sets* do not have size 2α for any α) yields the result that the universe also cannot have size 2α.

[9]Where this appeal to a metaphysically impossible antecedent is understood by appeal to facts about metaphysically impossible worlds.

by the hypotheticalist approach. This difference in treatment makes hypotheticalist approaches seem ad hoc and (ceterus paribus) unattractive.

Although fictionalist approaches differ from hypotheticalist approaches in that they take mathematical claims' surface logical form at face value, they face an analogous problem. It seems awkward, and violates the third desideratum, to say that ordinary people are engaging in make-believe when talking about numbers but not when talking about cities.

10.2.4 The Neo-Carnapian Approach

I will now use recent work on quantifier variance to articulate and develop a satisfying neo-Carnapian explanation of mathematicians' freedom to introduce new types of mathematical objects which avoids the problems above.

I propose that when mathematicians (or scientists or any other community) adopt coherent hypotheses characterising new types of objects, this choice can behave[10] like an act of stipulative definition, which simultaneously gives meaning to newly coined concepts and (slightly) changes the meaning of existence claims, as needed to make these hypotheses express truths – while preserving the usual inferential role of the existential quantifier ("$\exists$")[11] and the truth value of the vast majority of ordinary statements.[12] For example, when mathematicians introduce complex numbers by adopting certain claims about their relation to the real numbers, they shift the language we speak and change the meaning of our quantifiers to make sentences like, "there is a complex number which is the square root of -1" true. Similarly, sociologists' acceptance of ontologically inflationary conditionals like, "whenever there are people who ϕ, there is a tribe" can shift the language we speak and change the meaning of our quantifiers to ensure that this latter claim will express a truth.

According to my neo-Carnapian explanation, mathematicians' choices to start talking in terms of new objects (and certain explicit acts of stipulative definition) have the capacity to change the truth value of (rare) purely logical statements about the size of the universe like (the Fregean paraphrase of), "There are exactly 300,000

[10]The question of exactly what distinguishes adopting existence claims regarding a new kind of mathematical object as a mere hypothesis as opposed to as something like a stipulative definition is closely related to famously vexed issues about what it means to treat a claim as analytic. I won't say anything about this question here, since my aim in this paper is only to describe how genuine stipulative definitions can change our language and discuss how these can act as a model for the what mathematicians actually do in getting foundational knowledge of new mathematical structures.

[11]See Footnote 1.

[12]For example, stipulations introducing complex numbers should preserve the truth value of statements only quantifying over natural numbers or cats.

things."[13] Insofar as one takes the truth conditions for a sentence to be completely determined by the meaning (or sense, or contextual content) of its constituents, it follows that the meaning (or sense, or contextual content) of some piece of logical vocabulary which occurs within the Fregean paraphrase of, "there are fewer than 300,000 things" must change. It seems attractive to say that the meaning of "$\exists$" changes.

Taking this approach seriously requires that we can make sense of languages which 'change the meaning of the quantifier' in this fashion. This brings us to the topic of quantifier variance. In the next section, I will note that the form of quantifier variance which the neo-Carnapian explanation needs is weaker than the form of quantifier variance which has recently been used to argue for the defectiveness of ontology. However, even this version of quantifier variance is non-trivial, and I will have to defend it against the intelligibility worry mentioned at the beginning of this paper. First though, let me motivate interest in the neo-Carnapian story by noting how it promises to let us to explain mathematicians' freedom while satisfying the three desiderata above.

Unlike hypotheticalist approaches, the neo-Carnapian approach allows us to say that mathematical claims like, 'There is a number' and ordinary language claims like 'There is a city' use the same notion of existence and have directly analogous logical forms. The neo-Carnapian explanation merely posits that speakers' choice to accept some new kind of mathematical object can change exactly which notion of existence is relevant to that speaker's utterances.[14] Unlike fictionalist approaches, it takes both sentences above to express literal truths. Thus, it satisfies the third desideratum.

Unlike limitative approaches, the neo-Carnapian approach can accommodate all coherent choices of mathematical stipulations, without positing an arbitrary-looking joint in nature.[15] Thus, it also satisfies the first desideratum.

---

[13]Where, for example, the Fregean paraphrase of 'there are exactly two things' is '$(\exists x)(\exists y)[\neg x = y \wedge (\forall z)(z = x \vee z = y)]$'.

[14]Relatedly, one should note, that the above neo-Carnapian explanation of mathematicians' freedom does not require one to accept that normal English employs verbally different expressions corresponding to multiple different meanings that the existential quantifier could take on, e.g., to a metaphysically natural and demanding notion of existence and a laxer notion of existence. Thus, this view is not committed to endorsing statements like "composite objects exist, but they do not really exist." With regard to any context, we can fully agree with David Lewis that, "The several idioms of what we call 'existential' quantification are entirely synonymous and interchangeable. It does not matter whether you say 'some things are donkeys' or 'there are donkeys' or 'donkeys exist'... whether true or whether false all three statements stand or fall together."[13].

[15]One might worry that my approach removes appeal to arbitrary joints in nature re: the mathematical universe as a whole, but cannot remove the need to appeal to an arbitrary looking stopping joint in *set theory*. For, whether or not one takes the hierarchy of sets to contain instances for all possible structures available to be discussed by mathematics, one still has to suppose that this hierarchy stops at some point; it would be paradoxical to suppose that there are some objects, the ordinals, which contain segments corresponding to all possible ways for some objects to be well ordered (for this spine of ordinals would itself be well ordered, and by lemma 2.4 of [11] no

Finally, unlike social constructivist approaches, this view faces no pressure to claim that mathematicians' acts of stipulation somehow create or sustain mathematical objects' existence. Rather, we say that the choice to start talking in terms of new kinds of mathematical objects shifts us to a new language in which the relevant mathematical existence claims express a necessarily true proposition.[16] Thus, this theory has no problem satisfying the second desideratum. As different communities can speak different languages, this theory also has no trouble accommodating the idea that apparently incompatible claims about the total size of the universe can express truths in the mouths of speakers belonging to separated mathematical communities.

10.3 Quantifier Variance

Characterizations of quantifier variance in the current metaontology literature frequently combine two elements.

First, they include a multiplicity doctrine, QV_M. Someone who accepts this doctrine holds that the "∃" symbol can take on a range of variant meanings which are existential-quantifier-like, in that they satisfy the usual syntactic inference rules[17] associated with the existential quantifier. Additionally, they do not a recognize maximal sense of the quantifier, i.e., a quantifier meaning which we could view all these variant meanings of the quantifier as restrictions of.[18]

Articulations of Quantifier Variance in the metaontology literature also typically include a parity claim, QV_P, to the effect that all these meanings are (somehow) metaphysically on par. Although some meanings may be practically more or less

well ordering is isomorphic to an initial segment of itself). Thus, one might think that commitment to positing an arbitrary stopping point in mathematical reality cannot be permanently avoided. However, this is not so. I advocate a potentialist understanding of set theory which addresses the special Burali-Forti paradox concerning the height of the hierarchy of sets by eliminating the idea of a full set theoretic universe and using a mathematical trick to always interpret statements about the sets as statements about sets constructed before some V_α. While accepting the literal existence of ordinary mathematical objects like natural and real numbers and allowing that we could chose to talk in terms of sets going up to any ordinal height we can meaningfully characterize, I think we should respond to special and independently puzzling features of set theory which show up in the Burali-Forti paradox by denying that set theoretic claims really require the existence of suitable objects. See Putnam [15], Hellman[10] and [1] for details on this approach.

[16]See [6].

[17]Specifically, "(∃I) If $\Gamma \vdash \theta$, then $\Gamma \vdash \exists v\theta'$, where θ' is obtained from θ by substituting the variable v for zero or more occurrences of a term t, provided that (1) if t is a variable, then all of the replaced occurrences of t are free in θ, and (2) all of the substituted occurrences of v are free in θ'." and "(∃E) If $\Gamma_1 \vdash \exists v\theta$ and $\Gamma_2, \theta \vdash \phi$, then $\Gamma_1, \Gamma_2 \vdash \phi$, provided that v does not occur free in θ, nor in any member of Γ_2."[17].

[18]This definition is heavily influenced by Sider, e.g., by the discussion of quantifier variance in [19].

useful, on this view there's some deep sense in which none of them are better at capturing the nature of reality. Thus, for example, Chalmers characterizes Quantifier Variance as (roughly) the idea that, "there are many candidate meanings for the existential quantifier (or for quantifiers that behave like the existential quantifier in different communities), with none of them being objectively preferred to the other."[4] QV_P has been used to argue that many disputes in metaphysics and ontology are defective.

Although the above neo-Carnapian explanation of mathematicians' freedom is committed to QV_M, it is not committed to the more controversial metaphysical parity claim QV_P. Many philosophers on both sides of recent metaontological debates have been inclined allow non-metaphysicans substantial freedom to accept QV_M at least as it applies to non-metaphysicians. It is unsurprizing that anti-realists about ontology accept this claim. But even many realists accept it, holding that plumbers and scientists are free to employ a range of different notions of existence which are more ontologically profilgate than the single ontologically preferred one.[19] Thus, philosophers with a wide range of views on metaontology are in a position to accept my view.

10.4 Intelligibility Worries

Unfortunately, however, the neo-Carnapian proposal sketched above faces an immediate worry arising from its appeal to even the weak thesis QV_M. Are the range of alternate quantifier meanings which the neo-Carnapian explanation uses to explain mathematical stipulations really intelligible? In what sense would these variant notions be variant notions of existence?

Nearly everyone will allow that expressions like 'there is' can sometimes take on a *restricted* sense – as when someone says, "All the beers are in the fridge." However, many philosophers are inclined doubt the intelligibility of appeals to alternative quantifier-like senses for "∃" which are not mere restrictions of a fundamental, most generous, notion of existence which we use when doing ontology. For example, Wright and Hale claim not to understand "just what ... the postulated variant quantifier meanings [are] supposed to be."[21]. They maintain

[19]For example, Sider [19] has used quantifier variance (between the context of ordinary conversation and the metaphysics seminar) to capture the intuition that ordinary speakers, non-philosophical utterances like 'There's a hole in a sink' can express uncontroversially true statements, despite the fact that there's a deep open question about what exists in the more fundamental sense relevant to the metaphysics room. He says that there is a unique maximally natural sense of the quantifier which ontologists aim to employ, and that it is a deep open question whether holes exist in this sense. However, he allows that there is also a different (perhaps less than maximally ontologically insightful) sense which the quantifier can take on in the context of ordinary life/plumbing discussions, on which sentences like 'There is a hole in this pipe' can uncontroversially express truths.

that philosophers who appeal to quantifier variance owe an explanation of, "why the allegedly different quantifiers which can all be expressed by the words 'there are' are quantifiers and ... how they differ in meaning [from the quantifier currently in usage]."[21]. They grant that one can answer the first question by appealing to the existential quantifier's characteristic inferential role. However, they express grave doubts as to about whether the second question (what are variant existential quantifier meanings supposed to be like, and how do they differ from our own?) can be consistently and satisfyingly answered.

> The only obvious suggestion - that by introducing concepts of new kinds of objects (e.g. mereological sum, or number) we somehow enlarge the domain - is, in so far as it's clear, clearly hopeless. We cannot expand the range of our existing quantifiers by saying (or thinking) to ourselves: Henceforth, anything (any object) is to belong to the domain of our first-order quantifiers if it is an F (e.g., a mereological sum)'. For if Fs do not already lie within the range of the initial quantifier anything', no expansion can result, since the stipulation does not apply to them; while if they do, then again, no expansion can result, since they are already in the domain.[21].

In the absence of some consistent way of explaining what variant ontologically profligate quantifier meanings are supposed be like (and how they are to differ from our own), the neo-Carnapian explanation's appeal to these meanings can appear unacceptable and conceptually confused.

10.4.1 A Fregean Idea

In the rest of this paper, I will attempt to answer this challenge by describing how the alternate quantifier meanings invoked by the neo-Carnapian explanation function. I will do this by appealing to a broadly Fregean idea: that one can satisfactorily explain the meaning (or sense, or contextual content etc.) of logical operators like &, $\vee$ or $\exists$ merely by explaining how these operators systematically contribute to the truth or falsity of whole sentences in which they figure (and, perhaps, describing the characteristic inferential role associated with them). This idea is quite attractive, since it is hard to see how we could give any account of the meaning of these logical operators in any other way.

If one accepts this Fregean approach, it seems plausible that one can describe quantifier meanings associated with alternative languages (which hold the inferential role of the quantifiers fixed) in sufficient detail to answer Wright and Hale's challenge, just by systematically describing how truth conditions for sentences in these languages differ from those for sentences in our own language. For example, to the extent that we can understand truth conditions for assertions in our own language, such a systematic description allows us to systematically determine truth conditions for sentences in a new language which employs one of these alternate quantifier meanings.

Accordingly, I will answer Wright-and-Hale-style intelligibility worries about the kind of alternate quantifier meanings the neo-Carnapian explanation needs by giving a systematic account of truth conditions for utterances employing these alternate meanings. It will be important to remember that the aim of this story is entirely to (as Chalmers puts it), "[help us] understand the conditions under which various sorts of existence assertions are true or correct," so it need not, and should not, be understood as, "an account of the logical form of existence sentences [or] a conceptual analysis of these sentences."[4][20]

10.4.2 *Chalmers' Descriptive Strategy*

In 'Ontological Anti-Realism', David Chalmers describes alternative quantifier meanings (meanings which he takes to be precise sharpenings of the meaning of "there is" in ordinary English) by describing truth conditions for utterances employing these quantifier meanings. Chalmers' story provides a point of departure for my own approach. I will describe a simplified version of Chalmers' story, omitting various details which are irrelevant to the task at hand.

Chalmers describes variant precisifications of the ordinary English notion of existence which can vary the truth value of sentences like, 'There's a hole in this pipe' He does this by (in essence) associating each precise quantifier meaning with a function from possible worlds to set theoretic models.[21] This function associates each possible world w with a set theoretic model $f(w)$ which specifies both a domain representing how many objects are to count as existing (on this preciscification of the quantifier) at w, and extensions for all meaningful atomic relations and names used in ordinary English within this domain. Thus, for example, if (on some precisification of "there is"), it is true to say that there are holes, then the function associated with this notion of existence will associate the actual world with a set theoretic model which assigns the one place relation *hole*() a non-empty extension.[22] In this way, it assigns each possible world to a set theoretic model which

[20]If you prefer to say that uses of ∃ (with its usual inferential role) always has the same meaning (in the sense of Kaplanian character), but can differ in what content this determines in a particular context, Wright and Hale's challenge becomes a demand to explain the how content which (one claims) ∃ can on express in alternative contexts differs from the content which it expresses in our current context. The Fregean thought becomes the idea that one can adequately explain this difference by explaining how this contextual content systematically contributes to the truth conditions for whole sentences.

[21]Strictly speaking, Chalmers associates quantifier meanings with furnishing functions which take us from possible worlds to specifications of how some very specific (and otherwise undescribed) predicates $P_1 \ldots P_n$ apply and then takes our understanding of various English concepts to generate something like a set theoretic model from this furnishing function. However, see Footnote 19 in [4].

[22]Note that, not all mathematically possible furnishing functions will correspond to possible notions of existence, e.g., no sharpening of the English "there is" will correspond to a function which assigns a single object to the extension of both the relations *raven*() and *vegetable*() in the set theoretic model representing some possible world.

provides a "catalog of the objects which are taken to exist [at that] world." Note that the extension assigned to a relation like *hole*() at a given world will not be a set of holes but rather a set of abstract objects (in particular, other sets). In this way, we can describe quantifier meanings which make it true to say, "There are holes." without being committed to the existence of holes.

Appealing to these functions from possible worlds to models allows us to describe truth conditions for utterances employing alternative precisifications of the ordinary English "there is," as follows. Suppose one such precise meaning is associated with the a function from possible worlds to set theoretic models f. Then an utterance ϕ employing this meaning will be true at a possible world w if and only if it is true in the set theoretic model which f associates with that possible world. Thus, for example, an utterance of the form $\exists x F(x) \wedge G(x)$ will be true at exactly those possible worlds w such that $f(w)$ is a set theoretic model which assigns some object to the extension of both relations F and G.

Chalmers' descriptive strategy provides an attractive response to Wright and Hale's challenge to say what alternative quantifier-like meanings for "$\exists$" could be like. His story allows us to clearly and non-paradoxically describe alternate quantifier meanings which are not mere quantifier restrictions of our current sense. It also lets us explain why the various meanings for "$\exists$" so described would intuitively qualify as kinds of existential quantification, by appealing to the fact they retain their inferential role. For the standard definition of truth in a model ensures that any alternate meaning for "$\exists$", which is describable via the kind of function from possible worlds to set theoretic models mentioned above, will obey the standard inference rules.

10.4.3 The Problem

Unfortunately, however, there's a significant gap between the kind of quantifier meanings which Chalmers describes and those which the neo-Carnapian explanation of mathematicians' freedom requires. Accordingly, even if Wright and Hale accept Chalmers' construction as demonstrating the intelligibility of some ontologically profligate quantifier meanings, they might well persist in rejecting the kind of quantifier meanings required by the neo-Carnapian account. To defend the neo-Carnapian story we will need to describe the required quantifier meanings, and tell a story which goes beyond Chalmers' account in a few ways.

First, the neo-Carnapian story appeals to quantifier meanings associated with acts of stipulative definition that introduce entirely new (antecedently meaningless) mathematical vocabulary. Accordingly, we need to describe quantifier meanings associated with languages which recognize additional relation symbols (potentially expressing new concepts) as meaningful. As Chalmers' account works with a fixed list of meaningful relations it cannot describe truth conditions for such languages.

Second, there must be some plausible model of how acts of stipulative definition can switch a person between these different languages. Furthermore, the effects of these stipulations should be limited in scope. For example, stipulations introducing new mathematical sentences should not change the truth value of non-mathematical sentences like, 'The Nile is a river in Africa' or 'Carbon atoms have six protons.

Finally (as noted in the introduction), Chalmers' approach describes alternate quantifier meanings by associating possible worlds with set theoretic models, hence models whose domain is a set. For this reason, it can't describe quantifier meanings associated with languages which talk in terms of structures which involve proper class many objects – structures like the hierarchy of sets itself, or what you would get by stipulatively adding a layer of classes to the hierarchy of sets.[23,24] Thus, Chalmers' strategy seems incapable of describing exactly the kind of very ontologically proflegate alternative quantifier meanings which the neo-Carnapian strategy wants to invoke in order to say how it improves on limitative approaches like set theoretic foundationalism.

10.5 Structuralist Paraphrases

In the last two sections of this paper, I will propose a method for describing the kind of alternative quantifier meanings required by the neo-Carnapian explanation. My descriptive strategy differs from Chalmers' by eschewing appeal to any particular definite totality of objects (like the hierarchy of sets). Instead, it draws on a powerful, fundamentally modal, notion of 'logical possibility *given certain facts*'. I use this notion, which has arisen in the philosophy of mathematics literature,[25] to provide intended truth conditions for sentences employing an alternate meaning of the quantifier.

I will show how this descriptive strategy allows us to provide an attractive model of mathematical stipulation, one which honors the idea that mathematicians could stipulatively introduce mathematical structures too large to be identified with any portion of the mathematical universe recognized by our current language.

10.5.1 Logical Possibility

Many philosophers of mathematics have been inclined to acknowledge a notion of logical possibility which distinguishes 'coherent' descriptions of mathematical

[23]Perhaps appeal to nonstandard models would help (if one could prove some suitable analog to Skolem's theorem which applies to the appropriate languages for characterizing and talking in terms of new kinds of mathematical objects), but Chalmers does not seem to consider such models.

[24]Of course, if one took a potentialist approach to set theory, as discussed in Footnote 15, expressions like 'the structure of the hierarchy of sets' would not be meaningful so this worry would not arise.

[25]See, for example, [10] and [16]. The details of my presentation are most influenced by Hellman.

structures (like Hellman's second order Peano Arithmetic[26]) from 'incoherent' ones like naive set theory or Frege's theory of extensions[9]. This notion of logical possibility is, crucially, more generous (in that it counts strictly more scenarios as possible) than metaphysical possibility. It corresponds to a notion of logical necessity which is narrower than metaphysical necessity so that, to use Kit Fine's examples, "it is [logically] necessary that everything red is red, but not that nothing red is green or that I am a person"[8].

The notion of logical possibility resembles familiar notions of semantic consistency. However, facts about logical possibility are (importantly) taken to be primitive modal facts which do not to require grounding in the existence of 'witnessing' objects like set theoretic models, possible worlds or Tarskian re-interpretations of our language. When considering facts about logical possibility, we abstract away from metaphysically necessary constraints on the application of particular relations. Thus, for example, we will say that $\Diamond[(\exists x)Raven(x) \wedge Vegetable(x)]$ (where $\Diamond$ represents 'it is logically possible that...'), even if it would be metaphysically impossible for anything to be both a raven and a vegetable. We also abstract away from all constraints on the size of the universe,[27] so that $\Diamond[(\exists x)(\exists y)\neg x = y]$ is true regardless of how many things the actual world (or any metaphysically possible world) contains.[28]

10.5.2 Relative Logical Possibility

If one at accepts the intelligibility of this notion of logical possibility, it is only natural to also accept appeals to what is logically possible 'given' or 'relative' to certain facts. Consider a statement like the following.

> Given what cats and blankets there are, it is logically impossible that each cat slept on a different blanket last night.

This sentence has an intuitive reading which employs a notion of logical possibility *holding fixed the way that certain relations apply* (in this case, holding fixed what cats and blankets there are) rather than logical possibility *simpliciter*. A moment's thought will reveal that (on this reading) the above sentence is true if and only if there are more cats than blankets.

I propose to think of the logical possibility $\Diamond_{(...)}(\ldots)$ as an operator which takes a sentence ϕ and a finite (potentially empty) list of relation symbols $R_1, \ldots R_n$ and

[26] See [10]. This is essentially Peano arithmetic with the infinitely many instances of the induction schema replaced by a second order axiom of induction.

[27] See Etchemendy [7] on the tension between standard Tarskian reinterpretation-based accounts of logical possibility and the intuitive notion of logical possibility regarding this point.

[28] Thus, for example, even nominalist philosophers who take it to be metaphysically necessary that the universe can contain at most countably many objects can still acknowledge that it is logically possible for there to be uncountably many objects. See [14].

produces a sentence $\Diamond_{R_1,\ldots R_n}\phi$ which says that it is logically possible for ϕ to be true, without any change to how the relations $R_1, \ldots R_n$ apply. Thus, for example, the claim, 'Given what cats and baskets there are, it is logically impossible that each cat slept in a distinct basket' becomes:

C&B:$\neg \Diamond_{cat,basket} [(\forall x)(cat(x) \rightarrow (\exists y)(basket(y) \wedge sleptIn(x,y) \wedge (\forall z)[cat(z) \wedge sleptIn(z,y) \rightarrow x = z])]$

In making claims like the one above, we are discussing what is logically possible given the way that certain relations (in this case 'cat()' and basket()') apply. Note that, unlike pure logical possibility claims of the form $\Diamond\phi$ (which are always necessarily true or false), sentences involving the subscripted logical possibility operator $\Diamond_{(\ldots)}(\ldots)$ like C&B can be metaphysically contingent. Intuitively speaking, it is metaphysically possible for there to be so many cats and so few baskets that it would be logically impossible (given what cats and baskets there are in this scenario) for every cat to sleep in a different basket. Thus, it is metaphysically possible that C&B. But it is also metaphysically possible that ¬C&B.[29]

One can also nest claims about logical possibility, i.e., make claims about the logical possibility of scenarios which are themselves described using appeals to logical possibility, treating the above $\Diamond_{(\ldots)}(\ldots)$ operator as logical vocabulary like $\wedge, \vee, \neg, \exists$ and $\forall$. Doing this allows us to formulate and understand assertions like, "It is logically possible that it is logically impossible, given what cats and blankets there are, that each cat slept in a different blanket.", i.e., it is not incoherent to suppose that there aren't enough baskets for each cat to have slept in a different basket. More formally, this sentence becomes:

$\Diamond$**C&B**: $\Diamond(\neg \Diamond_{cat,basket} [(\forall x)(cat(x) \rightarrow (\exists y)(basket(y) \wedge sleptIn(x,y) \wedge (\forall z)[cat(z) \wedge sleptIn(z,y) \rightarrow x = z])])$

The above sentence, $\Diamond$(C&B), expresses a truth because (reading from the outside in):

- It is logically possible (holding fixed nothing) that there are 4 cats and 3 baskets.
- Relative to the logically possible (whether or not it's metaphysically possible) scenario where there are 4 cats and 3 baskets, it is not logically possible, given what cats and baskets there are, that each cat slept in a basket and no two cats slept in the same basket.

Finally, note that truth value of a sentence with the form $\Diamond_{R_1,\ldots,R_n}\phi$ does not depend on anything about relation symbols other than $R_1, \ldots, R_n$ appearing in

[29]If we accepted the existence of objects like David Lewis' possible worlds[12] (as I do not), we might express the above ideas as follows. Sentences involving the $\Diamond$ can be true at some metaphysically possible worlds and false at others. Just as the truth value of "$(\exists x)\text{cat}(x)$" at a possible world w will depend on the extension of 'cat()' at that possible world, the truth of "$\Diamond_{\text{cat}}[(\exists x)\text{cat}(x) \wedge (\exists y)\text{dog}(y)]$" at a possible world w will depend on the extension of 'cat()' at w (though not on the extension of 'dog()'). Both sentences above will be true at exactly those possible worlds where there is at a cat (because this is what it takes for it to be logically possible, given what cats there are, that there is a cat and a dog).

ϕ, such as their meaning or their extensions. This allows us to make intuitive sense of some $\Diamond_{R_1,\ldots,R_n}\phi$ sentences where ϕ involves relation symbols which lack any associated meaning/interpretation in our language. Thus, for example, there is intuitively a sense in which both, "$\Diamond_{cat}[(\exists x)(cat(x) \wedge happy(x) \wedge \neg happy(x))]$" and "$\Diamond_{cat}[(\exists x)(cat(x) \wedge foo(x) \wedge \neg foo(x))]$" express false propositions, while "$\Diamond_{cat}[(\exists x)(cat(x) \wedge happy(x))]$" and "$\Diamond_{cat}[(\exists x)(cat(x) \wedge foo(x))]$" both express truths. I will use sentences of this kind to help streamline the presentation of my story about the effects of stipulative definitions which introduce new vocabulary, below. However, if desired, one could reformulate my proposal (with a slight loss in simplicity/uniformity) to avoid such sentences, by using suitably chosen appeals to meaningful but otherwise un-mentioned relation symbols instead.

10.6 Putting It All Together

In *Mathematics Without Numbers*[10], Geoffrey Hellman draws our attention to approximately the notion of logical possibility discussed above, as part of an effort to provide hypotheticalist paraphrases for statements which appear to assert the existence of mathematical objects. According to Hellman, these hypotheticalist paraphrases reflect the true logical form of mathematicians' utterances. Saying this allows him to honor the intuitive idea that mathematicians assert truths, while remaining agnostic as to whether there are any abstract mathematical objects – for on his view, no ontological commitment to mathematical objects is incurred by accepting (ordinary) mathematical existence claims. Like other hypotheticalist accounts, Hellman's proposal suffers from the weakness that it doesn't treat mathematical and non-mathematical existence claims similarly.

However, nothing prevents us from availing ourselves of the power of logical possibility while avoiding the problems which beset hypotheticalist philosophies of mathematics. I will continue to make use of Hellman-style paraphrases as a tool for describing the truth conditions for certain mathematical utterances – in particular, mathematical statements in alternative languages associated with mathematical practices which talk in terms of different kinds of mathematical objects. However, I reject the idea these paraphrases reflect the true logical form of the propositions expressed by these utterances.

Instead I will say, as per the neo-Carnapian story described above, that claims like, "There are numbers." and, "There are cities." express (in our language) propositions which have the same logical form and employ the same notion of existence. After an act of stipulation which gets us to start talking in terms of new objects, this pair of sentences will continue to express propositions which have the same logical form and employ the same notion of existence (as one another). However, such an act of stipulation can change *which* single notion of existence occurs in the propositions expressed by both sentences. In this section, I will show how to use Hellman-style paraphrases to describe the alternate quantifier meanings brought into play by acts of mathematical stipulative definition. I will also provide an attractive model for the effects which acts of mathematical stipulation have.

As an example, let's consider how one might stipulatively introduce talk of the natural numbers into a 'nominalistic' language which does not talk in terms of any abstract objects.

For expository simplicity, I will discuss the effects of stipulations within a 'partially interpreted' language I_0 which contains all expressions which can be built up from finitely many meaningful relation symbols, the usual first order logical vocabulary[30] and the subscripted $\Diamond$, alongside arbitrarily many *currently unused* atomic relation symbols.

In this specific case, I will take I_0 to behave like a version of English which doesn't talk in terms of any mathematical objects. Thus, the meaningful (i.e., interpreted) relation symbols of I_0 will include expressions like "rat()" and "happy()", which apply in a way that idealizes the behavior of orthographicaly similar expressions in ordinary English, while other symbols like "ℕ()" and "foo()" will be among the meaningless relation symbols. I will further suppose that the objects acknowledged by I_0 are partitioned by a finite list of kind terms, in the following sense: there's a list of meaningful atomic relation symbols $K_1, \ldots K_n$ in I_0 such that, "$\forall x(K_1(x) \vee \ldots K_n(x)) \wedge \forall x[(K_1(x) \rightarrow \neg K_2(x) \wedge \ldots \neg K_n(x)) \wedge \ldots (K_n(x) \rightarrow \neg K_1(x) \wedge \ldots \neg K_{n-1}(x))]$" expresses a metaphysically necessary truth in I_0. For example, we might suppose that I_0 has a single expression 'physical object()', or a pair of expressions 'physical object()' and 'sociological object()' whose extensions necessarily cover all objects which unrestricted quantification in I_0 ranges over.

I will then define a special class of *nice* acts of stipulative definition, and describe the effects of making such stipulations. In doing this, I don't aim to capture every way that one might start speaking in terms of new kinds of objects, or even all possible acts of stipulative definition relevant to mathematics. I merely aim to provide a sufficiently plausible and concrete model for many kinds of mathematical stipulation to convince the reader of the attractiveness of the neo-Carnapian approach, and block the version of Wright and Hale's worries about the intelligibility of the quantifier senses associated with such acts of stipulation.

In general, I will think of acts of acts of stipulation as shifting a speaker from talking in their current language to talking in a closely related one. I will provide a simple translation procedure which pairs up sentences ϕ in the language I_S, which speakers of I_0 would start speaking if they made a nice stipulation S, with corresponding sentences ϕ' in I_0 that have the same truth conditions. I will then note how a nice act of stipulative definition could be used to introduce talk of numbers into the nominalistic language I_0 described above.

10.6.1 The Proposal

Consider a paradigmatic case of stipulative definition: one may try to introduce the property of bachelorhood by adopting the sentence, "$(\forall x)$(bachelor(x) $\leftrightarrow$

[30] $\exists, \forall, \wedge, \vee, \neg, \rightarrow$ (with its usual inferential roles)

man(x) $\wedge$ ¬married(x))" as something like an axiom, fixing the meaning of the previously undefined relation symbol 'bachelor()' in terms of the meaning of already understood relation symbols like 'married()' and 'man()'.

I will think of acts of stipulative definition as involving two elements. First, there is a list of relation symbols whose extensions this act of stipulative definition does not have to hold fixed.[31] I will call these the relation symbols being defined. In this case, 'bachelor()' is the only such relation symbol. Second, there is a sentence which one uses as something like an axiom to fix suitable new meanings for the relation symbols being (re) defined. In this case, the sentence is "$(\forall x)$(bachelor(x) $\leftrightarrow$ [man(x) $\wedge$ ¬married(x)])".[32]

I will say that an act of stipulative definition is **necessarily satisfiable** if it is metaphysically necessary that it is logically possible for the sentence being stipulated to be true, given the application of all meaningful relations which are not being explicitly re-defined. I will say that an act of stipulative definition is **necessarily categorical** if it is metaphysically necessary that it is not logically possible, given the facts about all relations not being explicitly re-defined, to satisfy this stipulation in two different non-isomorphic ways. See Appendix 2 for a precise definition of both notions. Finally, I will say that a stipulation is **nice** iff it is both necessarily satisfiable and necessarily categorical.

I propose that making any nice stipulative definition involving a sentence D while speaking I_0 would shift one to speaking a related language I_D with the same formal syntax as I_0 such that:

ϕ is true in I_D iff $\Box_{R_1 \ldots R_k}(D \rightarrow \phi)$ is true in I_0

where $R_1 \ldots R_k$ is the complete list of antecedently understood relations not explicitly being (re)defined by this stipulation.[33]

[31] I.e., relation symbols such that the stipulative definition is allowed to change the extension (if any) associated with that symbol.

[32] I introduce this distinction because two different acts of stipulative re-definition can involve adopting the same sentence, but give rise to markedly different effects on the truth conditions for sentences because they differ as to which terms in this sentence they are attempting to hold fixed vs. stipulatively redefine.

[33] If one already accepts counterfactuals with metaphysically impossible antecedents, it might seem more parsimonious to use counter-possible claims in place of appeals to logical possibility to specify truth conditions for languages like I_D, e.g., one might use counter-possible conditionals like, 'If there were objects satisfying NUMS then ϕ' to describe truth conditions for sentences in I_D. I resist this approach because it's controversial whether we have a cogent non-trivial grip on metaphysically impossible counterfactuals. Consider claims like, "If it weren't the case that $2 + 3 = 5$, then it would still be the case that $2 + 2 = 4$". It is far from clear what the closest worlds where $2 + 3$ isn't 5 would look like, or whether we can meaningfully appeal to the kind of closeness relation on impossible worlds which this sentence requires. It is also unclear how one could characterize intuitively acceptable stipulations given that even incoherent stipulations will be true at some counter-possible world.

10.6.2 Stipulatively Introducing the Numbers

Now let us focus on the task of introducing talk of numbers into our nominalistic language I_0. One can uniquely describe the intended structure of the natural numbers (how the relations $\mathbb{N}, S, +$ and $\times$, are supposed to apply) using nested appeals to logical possibility. I will call this description $PA_\diamond$ because it combines the standard first order Peano axioms (minus the instances of the induction schema) with a sentence $I_\diamond$, which expresses the same content as the second order induction axiom but uses the language of logical possibility and one otherwise unused relation symbol '$P()$'.[34]

$\mathrm{I}_\diamond$: $\Box_{\mathbb{N},S}[P(0) \wedge (\forall x)(\forall y)(P(x) \wedge S(x, y) \rightarrow P(y))] \rightarrow (\forall x)(\mathbb{N}(x) \rightarrow P(x))$[35]

This sentence says that, given the facts about what is a number and a successor, i.e., about how $\mathbb{N}$ and S apply, it would be logically impossible for P to apply to 0 and to the successor of each object which it applies to without applying to all the numbers. Thus $PA_\diamond$ has the following form:

$PA_\diamond = \mathbb{N}(0) \wedge (\forall x)(\mathbb{N}(x) \rightarrow (\exists y)(S(x, y) \wedge \mathbb{N}(y))) \ldots \wedge \mathrm{I}_\diamond$

This description $PA_\diamond$ uniquely pins down the intended structure of the numbers under plus, times and successor, in the following sense. Any two choices of an extension for $\mathbb{N}$ and ways for S, $+$ and $\times$ to apply to the objects within this extension would have to be isomorphic to one another.[36] However, stipulating $PA_\diamond$ as a definition of $\mathbb{N}$ and S does not qualify as a nice stipulation, i.e., it does not determine, for each metaphysically possible world, a unique way of modifying the universe (while holding fixed the extensions of all relations which are not being defined). For one thing, it does not determine how the relations $\mathbb{N}$ and S() are to apply to antecedently understood objects (for example, it is logically possible, given the way all antecedently understood relations apply, that $PA_\diamond \wedge (\exists x)[\mathbb{N}(x) \wedge \text{emperor(x)}]$). Additionally, requiring the truth of $PA_\diamond$ doesn't prevent one from adding extraneous new objects which are not numbers (or uniquely specify how many such objects one is supposed to add).

However, one can turn $PA_\diamond$ into a necessarily categorical stipulation by adding some extra clauses specifying how the numbers are supposed to relate to the antecedently recognized kinds terms $K_1 \ldots K_n$ (mentioned above) as follows.

NUMS: $PA_\diamond \wedge \forall x[K_1(x) \vee K_2(x) \ldots K_n(x) \vee \mathbb{N}(x)] \wedge \forall x[K_1(x) \vee K_2(x) \ldots K_n(x) \rightarrow \neg\mathbb{N}(x)] \wedge (\forall x)(\forall y)(\forall z)[(S(x, y) \rightarrow \mathbb{N}(x) \wedge \mathbb{N}(y)) \wedge (+(x, y, z) \vee \times(x, y, z) \rightarrow \mathbb{N}(x) \wedge \mathbb{N}(y) \wedge \mathbb{N}(z))]$

---

[34]Note that, we could have used any other one place relation symbol in I_0 (other than $\mathbb{N}$ and S) to state a version of the second order axiom of induction, e.g., we could have expressed the same constraint on how the numbers are supposed to be related by successor by instead saying $\Box_{\mathbb{N},S}[\text{happy}(0) \wedge (\forall x)(\forall y)(\text{happy}(x) \wedge S(x, y) \rightarrow \text{happy}(y))] \rightarrow (\forall x)(\mathbb{N}(x) \rightarrow \text{happy}(x))$

[35] Where a formula of the form $\psi(0)$ is shorthand for $(\exists z)(\forall w)\,(\mathbb{N}(z) \wedge \neg S(w, z) \wedge \psi(z))$.

[36]See Appendix 2 for how this notion can be expressed in terms of logical possibility.

In essence, NUMS says that there are numbers related to one another as per $PA_{\diamond}$, all non-numbers belong to one of the antecedently meaningful kinds $K_1 \ldots K_n$, there is no overlap between the numbers and these antecedently meaningful kinds $K_1 \ldots K_n$, and the relations S, $+$ and $\times$ only apply to numbers.

Now consider an act of stipulative definition which attempts to secure the truth of NUMS, and enjoys permission to define '$\mathbb{N}()$', 'S(,)', '+(,,)' and '×(, ,)'. This stipulation is necessarily satisfiable (as one can coherently add numbers to any metaphysically possible universe of physical objects) and necessarily categorical. Therefore, by the thesis above, making this stipulation in I_0 would shift one to speaking a related language I_{NUMS} such that:

ϕ is true in I_{NUMS} iff $\Box_{R_1 \ldots R_k}(NUMS \to \phi)$ is true in I_0

where $R_1 \ldots R_k$ is the list of antecedantly understood relations not allowed to be redefined by this stipulation (in this case, all the fintely many relation symbols which I_0 takes to be antecedently meaningful).

In the remainder of this section, I will discuss how claims in the new language introduced by this stipulative definition turn out to have intuitively correct truth values.

For one thing, this translation assigns correct truth conditions to all statements in the language of arithmetic (including those which cannot be decided via any proof procedure which we accept). Because *NUMS* includes a categorical description of the intended structure of the numbers under S, + and ×, for every sentence ϕ in the language of arithmetic, either $\Box(NUMS \to \phi)$ or $\Box(NUMS \to \neg\phi)$. Thus, we have either $\Box_{R_1 \ldots R_n}(NUMS \to \phi)$ or $\Box_{R_1 \ldots R_n}(NUMS \to \neg\phi)$. So the paraphrase indicated above does indeed ensure that for every sentence ϕ in the language of number theory either ϕ or $\neg\phi$ comes out true.

This translation also provides correct truth conditions for purely physical sentence like "$\exists x\, \mathrm{rat}(x)$." Suppose that there is at least one rat. My simple story correctly predicts that "$\exists x\, \mathrm{rat}(x)$" will express a truth in I_{NUMS}, as follows. Given what rats there are, it's logically necessary that $\exists x\, \mathrm{rat}(x)$. Accordingly, it's logically necessary, given the facts about how 'rat()' (and the various other relations $R_1..R_n$ which are not being stipulatively redefined) apply, that $\exists x\, \mathrm{rat}(x)$. Thus we have $\Box_{R_1 \ldots R_n}[NUMS \to \exists x\, \mathrm{rat}(x)]$ as desired.

An analogous story could be told about how we could use stipulative definition to introduce talk of mathematical structures satisfying some other categorical description D, e.g., to introduce the real numbers, or the hierarchy of sets up to any describable ordinal. It can also be used to capture how we could use stipulative definition to modify languages richer than I_0, and more like our own. See Appendix 3 for details.[37]

[37]We can also describe the effects of acts of stipulative definition on a language J_0 which contains a metaphysical possibility operator $\Diamond M$ and allows simple claims about metaphysical possibility in the obvious way.

10.6.3 Applied Mathematics

The above proposal also yields intuitively correct truth conditions for various statements of applied mathematics. For example, consider the claim that there are a prime number of rats. I will provide a natural example of how stipulatively introduced mathematical objects can figure in claims about non-mathematical objects, by formalizing this statement in I_{NUMS}.

The claim that there are a prime number of rats can be expressed in I_{NUMS} by saying that there is a prime number n such that it would be logically possible for some otherwise unmentioned relation $(Z(,))$ to bijectively map the rats to the numbers below n. With some abbreviations (including using $\phi(\cdot) \cong_Z \psi(\cdot)$ to abbreviate the claim that a two place relation Z **bijectively maps** the objects satisfying a formula $\phi(\cdot)$ to those satisfying $\psi(\cdot)$) we can express the claim that there are a prime number of rats as follows[38]:

$$\Diamond_{\mathbb{N},+,\times,rat}[(\exists x)(\mathbb{N}(x) \wedge \text{ prime}(x) \wedge [\text{rat}(\cdot) \approx_Z < (\cdot, x)])]$$

By the translation strategy discussed above, this sentence should express a truth in I_{NUMS} iff the following sentence expresses a truth in I_0

$$\Box_{R_1\ldots R_n}(NUMS \to \Diamond_{\mathbb{N},+,\times,rat}[(\exists x)(\mathbb{N}(x) \wedge \text{ prime}(x) \wedge [\text{rat}(\cdot) \cong_Z < (\cdot, x)])])$$

It is straightforward, if slightly involved, to verify that this claim is indeed true iff there are a prime number of rats. For example, suppose there are a prime number of rats. Then it is logically necessary that if there were numbers as per NUMS in addition to these rats, it would be logically possible to bijectively pair the numbers below some prime number n with the rats.

- If ϕ is a metaphysical-possibility-free sentence, $\Diamond_M\phi$ is true in J_1 iff $\Diamond_M[\Box_{R_1\ldots R_k}(D \to \phi)]$ is true in J_0.
- Truth functional connectives like $\wedge, \vee, \neg$ etc. contribute to truth conditions for whole sentences in the usual way, e.g., a sentence of the form $\phi \wedge \psi$ is true in J_1 iff ϕ is true in J_1 and ψ is true in J_1.

Thus, we can capture the effects of stipulative definitions made by speakers of a language J_0 which has sufficient expressive power to formulate the conditions for necessary satisfiability and necessary categoricity discussed above.

[38] Let us say that a two place relation Z **bijectively maps** the objects satisfying a formula $\phi(\cdot)$) to those satisfying $\psi(\cdot)$ (written $\phi(\cdot) \approx_Z \psi(\cdot)$) iff

- Z behaves like a function $\forall x \forall y[Z(x,y) \to (\forall z Z(x,z) \to y = z)]$
- Z is 1-1 $\forall x \forall y \forall z[Z(x,z) \wedge Z(y,z) \to x = y]$
- The domain of Z contains exactly the objects satisfying $\phi(x)$ $\forall x[\exists y Z(x,y) \leftrightarrow \phi(x)]$
- The range of Z contains exactly the objects satisfying ψ $\forall y[\exists x Z(x,y) \leftrightarrow \psi(y)]$

We further define prime(x) to abbreviate $(\forall y)(\forall z)[\times(y,z,x) \to (x = y \vee y = z)]$ and $< (\cdot, x)$ to abbreviate $(\exists k, k')S(k,k') \wedge +(\cdot, k', x)$

Thus, this account gives the right truth conditions for statements about both pure and applied mathematics.[39,40]

10.7 Conclusion

In this paper, I have shown how appeals to logical possibility can be used to give an attractive account of how acts of mathematical stipulation can be understood as changing our language (including the meaning of the existential quantifier). Like Chalmers' set theoretic proposal, the story above explains how certain alternate quantifier meanings differ from our own, by systematically describing truth conditions for assertions employing these meanings. It also lets us explain why these notions are quantifier-like, by noting that the standard inference rules for the existential and universal quantifiers will remain truth-preserving.

Unlike Chalmers' story, this account allows us to explain mathematicians' freedom to introduce new mathematical structures, including ones which are too large to have any (standard) models within the hierarchy of sets. This allows us to articulate an attractive neo-Carnapian explanation of mathematicians' freedom to introduce new objects: one which combines limitative approaches' parallel treatment of mathematical and non-mathematical existence claims with hypotheticalist approaches' avoidance of arbitrary limits on mathematicians' freedom. In so doing, it satisfies the three desiderata introduced at the beginning of this paper.

[39]Hartry Field has pointed out Hellman faces a problem about how to capture intended truth conditions for more complex statements of applied mathematics like, 'this backpack weighs 3.7 times more than that one', without appealing to infinitely many atomic predicates. This problem does not apply to us, insofar as we are using a language which talks in terms of mathematical objects (contemporary English) even when we use this language to describe the effects of stipulations which introduce additional mathematical objects. For, the technique above lets one capture stipulations which hold fixed the meaning of current mathematical vocabulary like 'realNumber()' and 'hasMassRatio(,,)'. Thus, if we are currently speaking a language which talks in terms of real numbers, and uses a relationship to the real numbers to measure ratios, there is no problem using my strategy to explain how making stipulations introducing new abstract objects would change the meaning of our quantifiers while preserving these facts.

[40]Intuitively, the above translation doesn't just preserve the truth value of claims, but also preserves facts about necessity and contingency. For example, purely mathematical claims in I_{NUMS} like, 'There are infinitely many primes' are associated with $\Box_{R_1 \ldots R_n}(NUMS \rightarrow \phi)$ sentences where ϕ (or the negation of ϕ) is a logically necessary consequence of *NUMS*. Thus, all such sentences are either metaphysically necessary or metaphysically impossible. In contrast, intuitively contingent statements in I_{NUMS} like, "There are a prime number of rats." are associated with propositions whose truth value depends on facts about what is logically possible given the application of 'rat()' (and various other antecedently meaningful relation symbols whose application is contingent).

This fact is what allowed us to extend the above story to an account of the effects of stipulation on a language like J_0 which contains a metaphysical possibility operator to yield correct truth conditions for claims like 'It is metaphysically necessary that there are numbers' or 'It it metaphysically possible that there are a prime number of rats' in footnote 37.

Appendix 1: Logical Possibility And Set Theory

This appendix will relate the truth conditions for many claims about logical possibility to more familiar notions from set theory (and modal metaphysics). The story I am about to provide cannot be used as a definition of logical possibility, because (among other things) it does not provide correct truth conditions for claims about logical possibility which involve collections too large to fit in the set theoretic universe, e.g., claims about what is logically possible given the facts about what sets there are. My aim in providing it is simply to use readers' existing familiarity with set theory to unambiguously explain how my notation for nested claims about logical possibility is supposed to work.

If we were willing to accept (and employ a metalanguage which quantifies over) metaphysically possible worlds and objects at these possible worlds (and make the further, false, assumption that each meaningful relation has a set sized extension at every possible world w containing only n-tupples of objects at w) we could use set theory to give truth conditions (in the sense of sets of possible worlds at which a given sentence is true) for the kind of claims about logical possibility which I have invoked above, as follows.

As usual, I will associate each sentence with a set of possible worlds at which that sentence is true. Intuitively speaking, the truth value of a logical possibility claim at a given possible world w is completely determined by 'structural facts' about the size of the domain of objects which our language recognizes as existing at this possible world, and the extensions of all relations within this domain. The truth value of all such sentences is completely determined by the kind of facts which are preserved by any set theoretic model which correctly captures the size of the universe and the extensions of relations within this universe, i.e., those facts which are invariant under model theoretic isomorphism.

Accordingly, I will specify whether a sentence ϕ is true at a possible world w in two steps. First, I will specify a particular set theoretic model (using set theory with ur-elements) which captures all relevant facts about w. Then, I will specify what it takes for a logical possibility sentence, $\Diamond\phi$ to be true relative to such a model (or, indeed, relative to any model specifying an extension for the appropriate relation symbols).

> For each possible world w, let **the set theoretic model associated with** w have as its domain the set of objects which exist at that world, and assign to the extension of each relation symbol R, the set of n-tupples of objects at w which stand in the relation which R names.

Now let us cash out the idea that a certain set theoretic model $\mathscr{M}$ corresponds to a scenario in which the sentence ϕ (which may itself contain instances of the $\Diamond$ claims) expresses a truth as follows.

A formula ψ is **true relative to a model $\mathscr{M}$ and a partial assignment function ρ which takes the free variables in ϕ to elements in the domain of $\mathscr{M}$**[41] just if:

- $\psi = R_n^k(x_1 \dots x_k)$ and $\mathscr{M} \vDash R_n^k(\rho(x_1), \dots, \rho(x_k))$.
- $\psi =$ '$x = y$' and $\rho(x) = \rho(y)$. etc
- $\psi = \neg\phi$ and ϕ is not true relative to $\mathscr{M}, \rho$.
- $\psi = \phi \wedge \psi$ and both ϕ and ψ are true relative to $\mathscr{M}, \rho$.
- $\psi = \phi \vee \psi$ and either ϕ or ψ are true relative to $\mathscr{M}, \rho$.
- $\psi = \exists x \phi(x)$ and there is an assignment ρ' which extends ρ by assigning a value to an additional variable v not in ϕ and $\phi[v/x]$ is true relative to $\mathscr{M}, \rho'$[42].
- $\psi = \Diamond_{R_1 \dots R_n}\phi$ where ϕ is a sentence , and there is another model $\mathscr{M}'$ which assigns the same tuples to the extensions of $R_1 \dots R_n$ as $\mathscr{M}$ and ϕ is true relative to $\mathscr{M}'$ and the empty assignment function ρ_0.

Finally, ϕ is true relative to a model $\mathscr{M}$ iff it is true relative to $\mathscr{M}$ and ρ_0. This suffices to define truth conditions for all sentences of the form $\Diamond_{R_1 \dots R_n}\phi$. One can then build up truth conditions for larger sentences using the standard recursion clauses associated with first order logical expressions, in the obvious way.

Appendix 2: Isomorphism and Categorical Stipulations

This appendix will show how we can use appeals to logical possibility to express claims about isomorphism, and precisify the notion of 'nice' stipulative definitions employed in section 10.6.

A two place relation Z **isomorphically maps** the Ps under $R_1 \dots R_m$ to the P's under $R'_1 \dots R'_m$ iff

- Z behaves like a function $(\forall x)(\forall y)[Z(x, y) \rightarrow ((\forall z)Z(x, z) \rightarrow y = z)]$
- Z is 1-1 $(\forall x)(\forall y)(\forall z)[Z(x, z) \wedge Z(y, z) \rightarrow x = y]$
- Z maps all of the Ps to P's $(\forall x)[P(x) \rightarrow \exists y(P'(y) \wedge Z(x, y))]$
- Z maps from the Ps onto all of the P's $\forall y[P'(y) \rightarrow \exists x(P(x) \wedge Z(x, y))]$
- Z 'respects' the way relations the $R_1 \dots R_m$ and $R'_1 \dots R'_m$ apply to the P and P' i.e., $(\forall \mathbf{x})(\forall \mathbf{y})[P(\mathbf{x}) \wedge P'(\mathbf{y}) \wedge Z(x_1, y_1) \dots \wedge Z(x_n, y_n) \rightarrow ([R_1(\mathbf{x}) \leftrightarrow R'_1(\mathbf{y})] \wedge \dots \wedge [R_m(\mathbf{x}) \leftrightarrow R'_m(\mathbf{y})]])]$

Let me abbreviate the claim that a two place relation Z isomorphically maps the Ps under $R_1, \dots R_m$ to the P's under $R'_1, \dots R'_m$ with the symbols $\langle P; R_1, \dots R_m \rangle \cong_Z \langle P'; R'_1, \dots R'_m \rangle$.

Now to say that the Ps under $R_1 \dots R_m$ are **isomorphic to** the P's under $R'_1 \dots R'_m$ is simply to say that such an isomorphic mapping is possible, given the way that $P, R_1, \dots R_m, P', R'_1, \dots R'_m$ apply, i.e.,

[41] Specifically: a partial function ρ from the collection of variables in the language of logical possibility to objects in $\mathscr{M}$, such that the domain of ρ is finite and includes (at least) all free variables in ψ

[42] As usual $\phi[v/x]$ substitutes v for x everywhere where x occurs free in ϕ

$$\Diamond_{P,R_1,\ldots R_m,P',R'_1,\ldots R'_m}(\langle P;R_1,\ldots R_m\rangle \cong_Z \langle P';R'_1,\ldots R'_m\rangle)$$

To avoid dealing with the complexities of ambiguity and imprecise meanings in Section 10.6, I focused on a special class of **nice** stipulative definitions which 'determine a unique way of modifying the universe of objects acknowledged by our current language' in every metaphysically possible scenario. We are now in a position to make this informal notion precise.

Speaking very roughly, we'd like to say that a stipulation that ϕ which attempts to give meaning to relation symbols $R_1,\ldots,Rn$ is categorical iff any two ways of modifying the number of objects which actually exist and fixing extensions for the new relations to be expressed by $R_1,\ldots,Rn$ so as to make ϕ true (without tampering with the actual extensions of the other antecedently meaningful atomic relation symbols $Q_1\ldots Q_m$) would have to be isomorphic. We could then say that a stipulation is **necessarily categorical** if it is metaphysically necessary that this stipulation is categorical. That is (in the language of possible worlds) if *for each metaphysically possible world w* there is (up to isomorphism) only one way to extend the collection of objects which our current language counts as standing in relations expressed by the symbols $Q_1\ldots Q_m$ at w so as to make ϕ true.

Unfortunately, it is difficult to translate this idea into a precise statement about logical possibility, as we can't simply assume there are independent objects representing all the possible ways of satisfying ϕ which we can assert are isomorphic. Unlike a traditional set theoretic model based approach, appeals to logical possibility only tell us that it is possible to satisfy ϕ given the extensions associated with $Q_1\ldots Q_m$; they don't ensure the existence of any objects which witness this fact.

To overcome this difficulty, I will describe a way to think of the extensions of two antecedently meaningless relation symbols, P and P', which are not mentioned in ϕ as specifying the domains of distinct models 'witnessing' two (potentially) different ways of modifying the universe by adding some objects to those objects which currently figure in the extensions associated with $Q_1\ldots Q_m$, and then choosing extensions for $R_1,\ldots,R_n$ in such a way as to make ϕ true. The models associated with P and P' will exactly agree on the extensions of $Q_1\ldots Q_m$ (so all objects which some Q_i applies to will satisfy both P and P') but they may differ with regard to how many additional objects they contain and the extensions they assign to $R_1\ldots R_n$

To formally develop this idea, I must first introduce some notation.

First, we will want to talk about a version of the sentence ϕ, $\phi \upharpoonright P$, which says that (a) all relations which are treated as meaningful by the sentence ϕ only apply to objects in P and (b) (so to speak) ϕ would be true if the objects which satisfy P constituted the entire universe. To express (a) we simply conjoin the claim that $\forall x_1\ldots x_k[R_i^k(x_1,\ldots x_k) \rightarrow P(x_1) \wedge \ldots P(x_k)]$ for each k-place relation R_i^k among the finitely many atomic relations $R_1\ldots R_n$ which are being newly defined with the claim that $\forall x_1\ldots x_k[Q_i^k(x_1,\ldots x_k) \rightarrow P(x_1) \wedge \ldots P(x_k)]$ for each k-place relation Q_i^k among the finitely many atomic relations $Q_1\ldots Q_m$ whose extension is being held fixed.

To express (b) we assert a version of ϕ which restricts all the quantification occurring outside of any $\Diamond$ in ϕ to the extension of P. Thus, for example, if ϕ were the sentence '$\exists x \, \text{cat}(x) \wedge \Diamond_{cat}[\exists x \, \text{cat}(x) \wedge \text{cat}(y) \wedge \neg x = y]$' we would add the sentence '$\exists x(P(x) \wedge \text{cat}(y)) \wedge \Diamond_{cat}[\exists x \, \text{cat} \wedge \text{cat}(y) \wedge \neg x = y]$'. Note that we don't need to modify any sub-sentences of ϕ occurring inside the $\Diamond$. For the truth of a $\Diamond$ claim depends on nothing about the actual world. And the truth of a $\Diamond_{K_1..K_n}$ claim in ϕ only depends on the extensions of $K_1 \ldots K_n$. Thus pretending that the entire universe only contains the objects which satisfy a predicate P which applies to all objects related by some such $K_1 \ldots K_n$ cannot change the truth value of any such sentence.

Note that, $\Diamond_{Q_1 \ldots Q_m}\phi \upharpoonright P$ will be true if and only if $\Diamond_{Q_1 \ldots Q_m}\phi$ is true. After all, if $\Diamond_{Q_1 \ldots Q_m}\phi$ is true then $\Diamond_{Q_1 \ldots Q_m}[\phi \wedge (\forall x)P(x)]$ will also be true, because P is not mentioned in ϕ (or among the antecedently meaningful relations $Q_1 \ldots Q_m$). From this it is easily deduced that $\Diamond_{Q_1 \ldots Q_m}\phi \upharpoonright P$ will be true. Conversely, similar considerations ensure that if $\Diamond_{Q_1 \ldots Q_m}\phi \upharpoonright P$ is false then so to is $\Diamond_{Q_1 \ldots Q_m}\phi$.

Next we want to define an analogous way of relativizing ϕ to P', while being careful to allow the universe associated with P' to 'witness' a different way of assigning extensions to various relations in $R_1 \ldots R_n$ from that associated with P. We can do this by defining $\phi \upharpoonright P'$ in the same way as $\phi \upharpoonright P$ but systematically replacing $R_1 \ldots R_n$ with a corresponding unused relation symbols of the same arity $R'_1 \ldots R'_n$.

Using these notions, we can say that a stipulative definition ϕ which attempts to fix the extensions for relations symbols $R_1 \ldots R_m$ is **categorical** iff:

$\Box_{Q_1 \ldots Q_m}[\phi \upharpoonright P \wedge \phi \upharpoonright P' \rightarrow$ the Ps under $R_1 \ldots R_n$ are isomorphic to the P's under $R'_1 \ldots R'_n]$

where $Q_1 \ldots Q_m$ are the finitely many other atomic relation symbols whose extension this stipulation is not permitted to modify.

Accordingly, if $\blacksquare$ is the metaphysical necessity operator, we can say that such a stipulative definition is **necessarily satisfiable** iff:

$\blacksquare(\Diamond_{Q_1 \ldots Q_m}\phi)$

And it is **necessarily categorical** iff:

$\blacksquare(\Box_{Q_1 \ldots Q_m}[\phi \upharpoonright P \wedge \phi \upharpoonright P' \rightarrow$ the Ps under $R_1 \ldots R_n$ are isomorphic to the P's under $R'_1 \ldots R'_n])$

It is **nice** if both conditions above apply.

Appendix 3: Capturing More

In this section, I will discuss how one can use the framework of nice ontologically inflationary stipulations discussed above to model the effects of a range of different choices to introduce new kinds of mathematical objects.

First, one can use the framework above to describe the effects of introducing the numbers into a language which already talks in terms of certain antecedently understood types of mathematical objects $M_1 \ldots M_k$ and includes some atomic expressions like 'mathematical object' or 'abstract object'. Because speakers of this kind of language will want to introduce the numbers *as a new kind of mathematical object* we will not model them as stipulating NUMS:

NUMS: $PA_{\diamond} \wedge \forall x[K_1(x) \vee K_2(x) \ldots K_n(x) \vee \mathbb{N}(x)] \wedge \forall x[K_1(x) \vee K_2(x) \ldots K_n(x) \rightarrow \neg\mathbb{N}(x)] \wedge (\forall x)(\forall y)(\forall z)[(S(x,y) \rightarrow \mathbb{N}(x) \wedge \mathbb{N}(y)) \wedge (+(x,y,z) \vee \times(x,y,z) \rightarrow \mathbb{N}(x) \wedge \mathbb{N}(y) \wedge \mathbb{N}(z))]$

but rather something like NUMS'.

NUMS': $PA_{\diamond} \wedge \forall x[K_1(x) \vee \ldots K_n(x) \vee \text{ mathematical object}(x)] \wedge \forall x[\text{mathematical object}(x) \leftrightarrow M_1(x) \vee \ldots M_k(x) \vee \mathbb{N}(x)] \wedge (\forall x)[M_1(x) \vee \ldots M_k(x) \rightarrow \neg\mathbb{N}(x)] \wedge (\forall x)(\forall y)(\forall z)[(S(x,y) \rightarrow \mathbb{N}(x) \wedge \mathbb{N}(y)) \wedge (+(x,y,z) \vee \times(x,y,z) \rightarrow \mathbb{N}(x) \wedge \mathbb{N}(y) \wedge \mathbb{N}(z))]$

In essence NUMS' says that: everything belongs to one of the antecedently recognized high level kind terms, something is a mathematical object iff it is either one of the antecedently recognized kinds of mathematical objects $M_1 \ldots M_k$ or it is a number, the numbers don't overlap with any of these antecedently recognized mathematical objects, the relations introduced alongside them don't apply to non-numbers,[43] and the numbers satisfy $PA_{\diamond}$. As above, this is a necessarily satisfiable and necessarily categorical stipulation which introduces the numbers and ensures that every number is a mathematical object.

A similar strategy can be used in cases where we want to ensure the truth of some sentence involving a relation (rather than preserving the extension of that relation). For example, we can describe how a stipulation which introduces new non-set objects to a language which talks in terms of set theory with ur-elements continues to ensure the truth of a sentence like 'every non-set is an element of some set'.

Appendix 4: Describing the New Language

In this paper, we have demonstrated that speakers of a language could consistently shift to speaking a new, more ontologically profligate, language, in a manner that explains mathematicians' ability to stipulatively introduce new objects. However, one might worry that one could never give, in one's current language, (as opposed to describe in a meta-language as we have done here) a satisfying description of the alternative quantifier meaning associated with the language one would come to speak by making an ontologically inflationary mathematical stipulation. Indeed, as one (plausibly) can't even describe truth conditions for all sentences in ones own language, it may well be that a full description of truth conditions for sentences

[43]Of course, if one is trying to model the introduction of new types of mathematical objects which are supposed to include and extend antecedently recognized mathematical objects (c.f. how the complex numbers are supposed to include and extend the real numbers) this clause will need to be suitably modified.

in this new language is impossible. However, just as one can nonetheless say a lot about truth in ones own language, I will show that one can also say a great deal about truth conditions in the language which would result from making a stipulative definition (while speaking one's current language).

Admittedly, I've only discussed how speakers of the nominalistic language I_0 described above can manifest their understanding of statements involving alternative quantifier senses by systematically making translations as described above. In particular, speakers of I_0 can not assert via any sentence in their language what the translation strategy would be if they were to adopt some stipulation only have dispositions to make the appropriate translations.

However, if we consider a L_0 corresponding to a (partially interpreted) language which talks in terms of numbers and Gödel codes for sentences and contains a limited truth predicate such as 'is a true sentence in the truth-predicate-free fragment of L_0',[44] then speakers of L_0 can systematically describe the truth conditions for a wide range of sentences in the language they would speak after making a simple and exhaustive stipulation involving some truth-predicate-free sentence D as follows.

> For all integers the number] $\ulcorner \phi \urcorner$ codes a truth in the truth-predicate-free part of J_1 iff the number $\ulcorner \Box_{R_1 \ldots R_k}(D \to \phi) \urcorner$ codes a truth in the truth-predicate-free part of J_0.

where $R_1 \ldots R_k$ is the list of antecedently understood relations not allowed to be redefined by this stipulation (i.e., all antecedently meaningful atomic relations).

Intuitively, this allows the speakers of L_0 to formulate a single unified description (in their own language) illuminating what the alternative quantifier meaning associated with the language resulting from stipulating the existence of some mathematical structure would be like. Indeed, for most mathematical stipulations, such as a stipulation introducing the complex numbers, this statement would describe all the *interesting* effects of making such a stipulation – even though it wouldn't describe the effects of the stipulation on uncommon sentences involving assertions of truth. For example, it would describe the effect of the stipulation on the sentence, "There is a square root of -1," though not on, "It is true that 'There is a square root of -1'." Indeed, with slightly more work even these sentences could be handled leaving only the truth conditions of truly troublesome statements involving recursive application of the truth predicate unspecified by this description.

References

1. Berry, S. Manuscript. Modal structuralism simplified.
2. Boolos, G. 1999. *Logic, logic and logic*. Cambridge: Harvard University Press.
3. Braun, David, "Indexicals", The Stanford Encyclopedia of Philosophy (Summer 2012 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/spr2015/entries/indexicals/.

[44]Because of the standard paradoxes involving truth predicates, one cannot hope to have a binary-valued truth predicate for the entirety of ones language.

4. Chalmers, D. 2009. Ontological anti-realism. In: *Metametaphysics: New essays on the foundations of ontology*, ed. R. Wasserman, D. Chalmers, and D. Manley. Oxford/New York: Oxford University Press.
5. Cole, J. 2013. Towards an institutional account of the objectivity, necessity, and atemporality of mathematics. *Philosophia Mathematica* 21(1): 9–36.
6. Einheuser, I. 2006. Counterconventional conditionals. *Philosophical Studies* 127(3): 459–482.
7. Etchemendy, J. 1990. *The concept of logical consequence*. Cambridge: Harvard University Press.
8. Fine, K. 2002. Varieties of necessity. In *Conceivability and possibility*, ed. T. Gendler Szabó, and J. Hawthorne. Oxford/New York: Oxford University Press.
9. Frege, G. 1980. *The foundations of arithmetic: A logico-mathematical enquiry into the concept of number*. Evanston: Northwestern University Press.
10. Hellman, G. 1994. *Mathematics without numbers*. Oxford/New York: Oxford University Press.
11. Jech, T. 2003. *Set theory: The third millennium edition, revised and expanded*. Berlin/New York: Springer.
12. Lewis, D. 1986. *On the plurality of worlds*. Oxford/New York: Blackwell.
13. Lewis, D. 1990. Noneism or allism? *Mind* 99(393): 23–31.
14. Parsons, C. 2007. *Mathematical thought and its objects*. New York: Cambridge University Press.
15. Putnam, H. 1979. *Mathematics, matter, and method*. Cambridge/New York: Cambridge University Press
16. Shapiro, S. 1997. *Philosophy of mathematics: Structure and ontology*. New York: Oxford University Press.
17. Shapiro, Stuart "Classical Logic", The Stanford Encyclopedia of Philosophy (winter 2013 Edition) Edward Zalta (ed.) http://plato.stanford.edu/archives/win2013/entries/logic-classical/
18. Shapiro, S., and C. Wright. 2006. All things indefinitely extensible. In *Absolute generality*, ed. A. Rayo, and G. Uzquiano. New York: Oxford University Press.
19. Sider, T. 2009. Ontological anti-realism. In *Metametaphysics: New essays on the foundations of ontology*, ed. R. Wasserman, D. Chalmers, and D. Manley. New York: Oxford University Press.
20. Uzquiano, G. 1996. The price of universality. *Philosophical Studies* 129(1): 137–169.
21. Wright, C., and Hale, B. 2009. The metaontology of abstraction. In *Metametaphysics: New essays on the foundations of ontology*, ed. R. Wasserman, D. Chalmers, and D. Manley. New York: Oxford University Press.

Chapter 11
"There Is an 'Is' in 'There Is'": Meinongian Quantification and Existence

Francesco Berto

Abstract Against the mainstream Quinean meta-ontology, Meinongians claim: "There are things that do not exist". It is sometimes said that the "there are" in that sentence expresses "Meinongian quantification". I consider two supposedly knock-down meta-ontological objections to Meinongianism from the literature: (1) an objection from equivocation, to the effect that the view displays a conceptual or semantic misunderstanding, probably of quantificational expressions; and (2) an objection from analyticity, to the effect that sentence is Frege-analytically false i.e., it is synonymous with a logical falsity. Objection (1) is countered via a development of Williamson's argument against epistemic conceptions of analyticity. Objection (2), which points at alleged linguistic evidence, is countered by resorting to linguistic counter-evidence. The upshot is a set-up of the debate between Quineans and Meinongians, in which the two parties disagree on substantive matters concerning *de re* the property of existence, taken as a natural property in the Lewis-Sider sense; and in which quick alleged refutations, such as objections from meaning-variance or analytic falsehood, rarely achieve their expected results.

11.1 Two Fast Ways for Meinongianism to Go

Meinongians claim:

(M) There are things that do not exist.

The "there are" in (M) is sometimes said to express "Meinongian quantification" (see e.g. Lycan [15], Lewis [14], Burgess and Rosen [3], van Inwagen [36]). On the other hand, the mainstream reply to what has been called "the question of ontology" is:

F. Berto (✉)
Department of Philosophy and Institute for Logic, Language and Computation (ILLC), University of Amsterdam, Oude Turfmarkt 141, Office 2.17, 1012 GC Amsterdam, The Netherlands
e-mail: F.Berto@uva.nl

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_11

(Q) Everything exists.

(The "Q", as you may have guessed, reminds us of Quine). Or rather, that is the apparently *easy* reply to that question. Thus begins an introduction to ontology:

> It is customary to identify ontology with that branch of philosophy that originates from the question: "What exists?". And it is customary to specify that this question has two kinds of answer. The first answer is easy, if not trivial, and can be summed up in one word: "Everything". As Quine [[22]: 3] has written, everything exists because it makes no sense to speak of "nonexistent entities", and those who think otherwise would manifest, not an ontological disagreement, but a misunderstanding of the very concept of existence. Of course, one will say, elephants exist but not unicorns, nor round squares; this does not mean, though, that unicorns and round squares are things that do not exist. It just means that there are no such things. Precisely because it would be inconsistent to claim that *something does not exist*, though, to claim that *everything* exists is tautological, that is, devoid of content, therefore of interest. (Varzi [39]: 3)[1]

I will consider two objections to Meinongianism, nicely mingled in this quote. To be sure, they are not in there in full-fledged form; but they can be developed starting from claims included in it:

1. "It makes no sense to speak of 'nonexistent entities', and those who think otherwise would manifest, not an ontological disagreement, but a misunderstanding of the very concept of existence." Call the corresponding objection, the *objection from equivocation*. It can be developed as follows:

   On the face of it, (Q) and (M) appear to express contradictory propositions, thus engendering, when asserted by the Quinean and Meinongian respectively, an "ontological disagreement" – but they do not. The Meinongian asserting (M) as a motto summarizing her theory displays some basic misunderstanding: to use a Quinean catchphrase, she is "changing the subject" of some of the words she uses. Empirical psychology may be interested in investigating why the Meinongian lacks the linguistic or conceptual competence to see the point; but as far as ontology is concerned, Meinongianism is flawed for this reason.

2. It is "easy, if not trivial" to see that (Q) is true – and it is "easy, if not trivial" to see that (M) is false. Indeed, "to claim that *everything exists* is tautological", just as "it would be inconsistent to claim that *something does not exist*". Call the corresponding objection, the *objection from analyticity*. It can be developed as follows:

[1] My translation. Once the easy answer is given, of course, not everything is settled: "there remains room for disagreement over cases" (Quine [22]: 32).

(Q) and (M) really express contradictory propositions in the debate – thus, they cannot both be true. And (Q) expresses an easily recognizable analytic truth – precisely, a *Frege-analytic* one: one that is synonymous with a logical truth. Then (M) expresses an easily recognizable Frege-analytic falsity. (M) is, in fact, only one replacement of synonyms away from logical inconsistency. Meinongianism is a (nearly) immediately self-refuting position, and it is flawed for this reason.

The debate between Quineans and Meinongians has a long history, and most of its salient episodes antedate the recent, burgeoning development of *meta-ontology*. As far as I know, the term was introduced by Peter van Inwagen's [34] paper bearing that title. He wrote that meta-ontology deals with "questions about the meaning of being and questions about the proper method of [. . .] ontology" (van Inwagen [34]: 3). Meta-ontological questions have to do "with the intension, as opposed to the extension, of 'being'" (Ibid.). It is controversial whether intensions are sufficiently fine-grained for meanings. But the jargon of intension and extension nicely captures the idea of a higher level of difficulty: given that intensions determine extensions, a dispute on intensions-or-meanings is more fundamental than one on extensions.

The debate between Quineans and Meinongians is largely meta-ontological. A realist and a nominalist Quinean may give opposite replies to the question: "Do propositions exist?"; but at least they agree on what "exists" stands for here. With a Meinongian, not even such an agreement is guaranteed. My justification for dealing with the two objections above in one paper is that they are both naturally classified as meta-ontological. They call into play notions variously related to meaning, such as synonymy, analyticity, contradictoriness, (mis-)understanding. As we have begun to see, according to many Quineans (M) is *obviously* mistaken, or borders on the unintelligible:

> I really cannot understand Relentlessly Meinongian quantification at all; to me it is literally gibberish or mere noise. (Lycan [15]: 290)

> In sum, there are no things that do not exist. This thesis seems to me so obvious that I have difficulty in seeing how to argue for it. (van Inwagen [34]: 16)

> Surely there are no non-existent objects; surely that is a truism if anything is. (Stanley [30]: 39)

That the Meinongian has got some meaning wrong is a cover-all verdict on her situation, shared by (1) and (2). The two objections disagree, though, on what, exactly, has gone wrong with (her asserting) (M). In fact, they can hardly be raised together, for (2) assumes something that (1) denies, that is, that (M) and (Q) really express contradictory propositions in the debate. The two objections are different enough to trigger different responses.

11.2 Equivocation

The reply to objection (1) expands a line of thought due to Tim Williamson's methodological *The Philosophy of Philosophy*. So-called "epistemological conceptions of analyticity" (Boghossian [2]) take some selected sentences as such that understanding them is sufficient for taking them as true (believing them, being disposed to assent to them, etc.: fine-tuning the kind of acceptance is of minor importance here). Then failure to take them as true (etc.) is sufficient for misunderstanding. The objection from equivocation has it that (Q) is one such sentence. The Meinongian asserting (M) aims at contradicting (Q), but displays a misunderstanding: she lacks some linguistic or conceptual competence needed to fully grasp what she tries to deny. Then what (M) means in her mouth is not really the contradictory of what (Q) means (one may further wonder whether (M) means anything at all, or it is Lycan's "gibberish, or mere noise"; but change from meaningfulness to meaninglessness is meaning change already, and enough to fail to contradict).

On which word(s) is the Meinongian equivocating? This cannot be quickly settled by looking at whatever (Q) and (M) are explicitly about, where what a sentence "is about" is what (in the relevant context) its constituents refer to. Neither (Q) nor (M) are explicitly about language: they don't openly speak of linguistic items in order to ascribe meanings.

A driving force behind the development of meta-ontology has been the feeling of uneasiness induced in some by contemporary ontological debates. Such debates have usually been conducted in a (Q)-friendly environment: a broadly Quinean framework in which ontological questions are taken as quantificational questions. The quantifiers have been the obvious (or not so obvious, see Sider [28]: 387–91) suspects in the charge of meaning-change raised by some deflationary philosophers. According to Eli Hirsch [10], when two ontologists debate thus:

DKL: "There exist tables and coins-Eiffel tower fusions."
PVI: "No, there exist neither coins-Eiffel tower fusions, nor tables (only simples arranged table-wise)."

…The debate is shallow, i.e., involving no disagreement about worldly facts, because of equivocation in "there exist" between the two parties' claims (the twofold example embellishes the one in Sider [28], ibid; you may have guessed which ontologists "DKL" and "PVI" remind us of). Because of such equivocation, the two parties end up talking past each other by failing to express contradictory propositions. Perhaps Quineans and Meinongians, too, mean different things by the respective "there are" and "every" in (Q) and (M). Talk of "Meinongian quantification" may encourage this thought: the Meinongian has a deviant (if not flatly absurd) conception of the quantifier, because of her failing to come to grips with the meaning of the relevant quantificational devices of ordinary English.[2]

This sounds quite implausible. I agree with a committed Quinean who claimed: "the neo-Meinongians and I mean the same thing by the unrestricted quantifiers" (van Inwagen [37]: 53; I will soon come to the qualification "unrestricted"). For "every" and "there is" (and "some", and "any", and "most", and "all", etc.) are words a competent English speaker effortlessly uses in everyday talk. According to a widespread view (see Marconi [16]), competence in a given language is holistically entailed by full participation in the communication practices of the relevant linguistic community. Such a view may be neutral between externalist (*à la* Putnam-Burge), internalist, or use-theoretic perspectives on meaning. As Williamson stresses in his campaign against epistemic analyticity, it is vital for mainstream accounts of reference that the intention to use an expression with the referent it has in the community be normally successful. Meinongians are clear that they want their words to be interpreted as words of ordinary English, and this should be taken into account when assessing semantic competence. By this picture, native English-speaking Meinongians were linguistically competent before learning anything about ontology. They may have been led (or misled) into believing in (M) by generalizing from a large number of common-sense quantified claims they heard, since they were children, uttered by people from their linguistic community – such as, for instance:

> There is something which has been sought by many, namely the site of Atlantis, but it does not exist.
>
> I thought of something I would like to give you as a Christmas gift, but I couldn't buy it for you because it doesn't exist.
>
> Some of the things you're talking about don't exist.
>
> Some of the gods are tempestuous, but of course no gods exist.[3]

It is implausible that the quantifier changed its meaning in the Meinongians' mouth, or that they started – unbeknownst to them – to equivocate, when they moved from accepting and asserting sentences such as these to accepting and asserting (M). If native English-speaking Meinongians were linguistically competent at the time of their understanding and use of quantified sentences like the above, they

[2] To be sure, for authors like Hirsch both PVI and DKL can make a true claim given what each means by the respective quantifier. The Quinean raising the objection from equivocation will not normally grant such symmetry to the Meinongian. In her view, the Meinongian is merely making semantically deviant claims, or pseudo-claims. This may not be the view of the typical deflationist. I am not aware of deflationists explicitly addressing the issue of Meinongian quantification in the literature, but perhaps a Hirschean may take the (Q) vs. (M) debate as shallow, just as the PVI vs. DKL debate.

[3] The first example comes from Wolstertorff [42]; the second from Priest [20]; the last two from McGinn [17].

still are. Then "we cannot understand them better if we translate their word [of quantification] by some non-homophonic expression, or treat it as untranslatable".[4]

If this holds for the quantifiers, it also holds for the other bits of language involved in (Q) and (M). These include no technical philosophical vocabulary (neither sentence says anything about metaphysical necessities, ontological dependencies, the *a priori* knowability of something, etc.). They include, besides the quantificational devices, only everyday words like "thing", "not", and that predicate, "exist(s)". These too are mastered by competent speakers of English, and the point applies to them, too. It would be implausible to claim, for instance, that the Meinongian is deviant in what she means by "thing" – that she has some conceptual or semantic deficiency concerning *thinghood*.[5] The Meinongian could easily claim of "thing" what the aforementioned committed Quinean claimed of the cognate word "object": "an object is anything that can be the value of a variable, that is, *anything* we can talk about using pronouns, that is, anything" (van Inwagen [2002]: 180). So taken, "object" or "thing" are just unrestricted, maximally general terms: "Every x is such that, if x is a thing, then x is P" just means "Every x is such that x is P"; and "Some x is such that x is a thing and x is P" just means "Some x is such that x is P".

One may resist the claim that (Q) and (M) make for everyday talk. Unlike the quantified sentences displayed above, they are not the kind of claim the layman would easily make. On the contrary, they are likely to be heard almost only in the ontology room. What makes of both (Q) and (M) distinctively (meta-)ontological (contrast ordinary) claims, then, since it's not any of the words they are made of nor, arguably, their mode of composition, which is also pretty ordinary? It is the fact, I think, that the Quineans and Meinongians uttering them normally take their quantifiers as unrestricted. This may be the contextual effect of the ontology room: both parties want to talk about everything; they want what they claim to leave nothing out. They can agree on what they are doing as philosophers: they are engaging in the study of things qua things.[6]

Absolutely unrestricted quantification may be problematic.[7] Suppose we embrace what has been called, following (a certain interpretation of) some Cantorian claims, the Domain Principle: a meaningful quantified sentence presupposes a

---

[4]Williamson [41]: 91. Williamson's imaginary characters, Peter and Stephen, are taken as deviant in their use of the quantifier in their denials that every vixen is a vixen. Lewis [14] proposed to have Meinongianism collapse via non-homophonic translation into a generous form of Platonism, followed in this by Burgess and Rosen [3].

[5]Some interesting points concerning the ontologists? use of "thing" and "object" are raised by Thomasson [32].

[6]They may disagree on how to call such a study. Some Meinongian may resist the term "ontology", on the ground that it misleadingly injects being in the notion of thing, or object. She may prefer Gegenstandstheorie, "object theory". *This*, I take it, would be a merely terminological disagreement: it would only concern whether or not to use a technical term to label a certain philosophical sub-discipline.

[7]Thanks to Tuomas Tahko for pressing me on this point. A collection of essays investigating the subject of absolutely unrestricted quantification is Rayo and Uzquiano [23].

domain for its quantifiers to range over (see Priest [19], Part III). Then absolutely unrestricted quantification will bring with it the paradoxes of total sets. Thus such philosophers as Dummett have rejected absolutely unrestricted quantification on the basis of there being no universal set. I find the idea of absolutely unrestricted quantification plausible and intuitive. This just comes with the notion of everything without limitation, that is, of all things *qua* things. It is an idea we appeal to, it seems to me, all the time in philosophy, when we claim that everything is self-identical, or that everything is either an abstract object or a concrete one, or that everything complies with this or that law of logic: see Williamson [40] and van Inwagen [37] for (long and short, respectively) defenses of this view.

The issue is neither here nor there in the (Q) vs. (M) debate anyway. *Whatever* (contextual, minimal, domain-increasing) restriction has to be imposed on quantification, it applies to both parties' claims. Once the point is raised, both can make explicit that they want their quantifiers to be interpreted in the least restrictive way allowed by the right domain-increasing condition or limitation of generality. Their intentions again ought to be taken with Williamsonian seriousness.

This does not entail that the quantifiers are not involved in the substantive Quineans vs. Meinongians debate; to the extent that they are, talk of “Meinongian quantification” seems to me still appropriate. But their involvement depends on their being connected, or rather not, with the only remaining item at issue between (Q) and (M), that predicate, “exists”. Our paradigmatic Meinongian ought not to be interpreted as semantically deficient with respect to that predicate either. Again, she may have been (mis)led to endorsing the generalization (M) consists in by considering common-sense sentences like the ones above, which happen to include that predicate too. Philosophy-proof people claim, in all seriousness, that the Everest exists but Mount Doom does not; that horses exist, but not winged horses; that Troy has almost certainly existed but the same cannot be said of Atlantis; etc. The Meinongian's patterns of assent with respect to claims such as these are likely to manifest no semantic deviance.

However, focusing on “exists” can now help to get a better picture of what is going on in the (Q) vs. (M) debate. I agree with the aforementioned Quinean on how to give this better picture:

> You will misunderstand what I have been saying if you take me to have been saying that neo-Meinongians (on the one hand) and I (on the other) mean two different things by “exist”. The neo-Meinongians and I have different *theories* about what “exist” means [...]. When they use the English word “exist”, *they* mean by it what it means, and if that happens to be, as I say it is, “not-all-not”, they mean “not-all-not” by “exist” – although, according to their mistaken theory about the meaning of “exists”, that is not what they mean by it. (Van Inwagen [37]: 53)

I think this is exactly the right way to understand the debate. Giving a theory of existence is difficult, hence the long-standing disputes between radically opposite theorists. The layperson who has never been asked what existence is, though, is able to use “exists” in a wholly adequate way. There is no reason to expect her to lose such ability by her coming to develop a non-mainstream philosophical theory of existence.

11.3 Resetting the Debate

From now on I will confront Meinongianism and Quineanism, as summarized by (Q) and (M), as two opposite theories of the property of existence, i.e., of what the predicate "exists" refers to (one may immediately object that to phrase the debate in terms of the *property* of existence pre-judges the issue, to some extent, against the Quinean, but we shall see how this need not be the case). Many exponents of the two parties may also agree in taking existence as one of those properties Lewis would call (*perfectly*) *natural*. Another committed Quinean, namely Sider [28], has pursued such a route while arguing against ontological deflationism *à la* Hirsch. For Lewis, naturalness is a feature some properties have due to their marking objective, structural, mind-independent similarities in the world. Naturalness for a property comes in degrees, and (the more) natural properties make for (the greater) intrinsic similarity of things sharing them. Lewis believes the distinction between natural and non-natural properties to be a primitive. We can just give examples: greenness is much more natural than grueness because green things are similar in ways in which grue things cannot be, etc. (see Lewis [13]: 62–3).

In the Lewisian view, naturalness makes some properties work not only as determiners of similarity, but also of content: they act as "reference magnets", as semantic values for predicates, by determining reference above and beyond issues of language use and meaning-change. According to Sider, the charge of meaning-change raised by ontological deflationists can be countered by taking the property of existence as a (perfectly) natural reference magnet, intrinsically eligible for what "exists" is to stand for. Ontological questions are not shallow, insofar as they are substantive, structural questions about the nature of such property.[8]

Suppose the same considerations apply against our first objection. Then Meinongianism in its various versions can be taken as a family of theories giving characterizations – in fact, sometimes quite different from each other – of that natural property.[9] Quineanism in its turn consists of a family of theories giving alternative characterizations – again, sometimes quite different from each other –

---

[8]To be sure, for Sider such a structure making for the naturalness of existence is quantificational structure. But then, Sider is a Quinean.

[9]For instance, some Meinongians, e.g., Routley [26], Priest [20], characterize existence univocally as the having of causal powers, and/or spatiotemporal location. Others, and probably Meinong himself, have a more pluralistic approach. For Meinong there are two modes of being: existence properly so called (*Existenz*) for *concreta* and subsistence (*Bestand*) for *abstracta* (he did not talk this way, but this seems to me a fair reconstruction of his view in contemporary ontological terms). Things like Plato, Holmes, or Obama may (concretely) exist or not, whereas things like sets and functions may (exist the sense of) subsist or not. Existence-as-subsistence may be something like being consistent, or coherent, or well-defined for the involved notion. In this sense the mathematician claims that the set of integers and the operation of division by seven exist, whereas the Russell set and division by zero do not. For a classic introduction to Meinong's philosophy, see Grossmann [8].

of that natural property.[10] How the two families diverge can be summarized by pairing their answers to three questions about the property at issue. Is it a “first-order” property of things? (Quinean: Yes; Meinongian: Yes). Is it a blanket property, one that everything has? (Quinean: Yes; Meinongian: No). Is it definable via logical notions? (Quinean: Yes; Meinongian: No).

Having the Quinean reply “Yes” to the first question may seem odd.[11] She will typically place herself in the tradition stretching back to the Kantian claim: “‘*Being*’ is obviously not a real predicate; that is, it is not a concept of something which could be added to the concept of a thing” (Kant [12]: 505). Kant opposed *real* or *determining* predicates to *merely logical* ones, and he took as logical those predicates that add no information on the referent of the subject term, no matter what this is: they stand for trivial features of anything whatsoever. Good examples might be “is blue if blue”, or “is either a chair or not a chair”.

However, according to authors like Gareth Evans or Nathan Salmon the Quinean, too, ought to grant that existence is a “first-order” property of things, though not a real one in the Kantian sense. The key meta-ontological thesis of Quineanism is that “Existence is adequately captured by the existential quantifier” (van Inwagen [34]: 18).[12] Then we can define the relevant existence predicate, in the familiar way, as $Ex =_{df} \exists y(y = x)$: to exist is to be (identical with) something. From the *definiens* we obtain an expression that, Salmon claims, plainly denotes the property at issue. Technically, we may use a lambda-operator, and abstract $\lambda x.\exists y(y = x)$, “the property of being an x, such that x is (identical with) something”. Now:

> Each of the notions involved in the definition of the predicate “exists” is precise and mathematically respectable; each of the expressions making up the definiens has a definite sense or content. In fact, each of the three notions involved – existential quantification, identity, and abstraction – is precise in a way that many everyday notions are not. [...] If a set of expressions that express concepts or attributes as their sense or content are appropriately combined to form a new expression, the compound expression thus formed has a sense or content that is determined in a certain way by the senses or contents of the combined component expressions. Hence the phrase “is identical with something”, and the displayed expression, express a definite property or concept as their (shared) sense or content. This is the property or concept of being identical with something (or more simply,

[10]For instance, some philosophers (e.g. McGinn [17]) seem to conflate the Quinean meta-ontological view of existence as quantification with the broadly Fregean-Russellian view that existence is reducible to a higher-order feature of some abstract objects (Fregean concepts, properties, or Russellian propositional functions): that of being instantiated. But van Inwagen [37] points at some plausible differences between the two traditions. Though I cannot argue it here, I believe such differences not to prevent a uniform assessment of the two views at a meta-ontological level.

[11]Thanks again to Tuomas Tahko for pressing me on this point. I hope I have adequately addressed it in what follows.

[12]Van Inwagen adds: “... of formal logic”. But later on, he makes clear that “the meaning of the [formal] quantifiers is given by the phrases of English – or some other natural language – that they abbreviate” (Ibid.). The notation of “formal logic” just helps to clarify the logical structure of quantified sentences of ordinary English.

> the property or concept of being something). It is this property or concept that is the sense or content of the predicate "exists". (Salmon [27]: 64)[13]

. . . And this is a "first order" property of things – what the variable *x* can take as values in the formulas above. The point of Quineanism is that the property at issue is reduced to the quantifier – and identity. Both quantification and identity are logical notions. So existence is, in a precise sense, a logical property. And no doubt, it is a blanket property. If to exist is nothing more than to be something (a "thin conception of being", van Inwagen [35] calls it), since everything is something, you ought to assert (Q). For the Meinongian, on the other hand, "exists" is a predicate standing for a real property of things: one which, as per (M), not all things possess, and which cannot be reduced to other notions – in particular, logical ones – through definition.

11.4 Theoretical Equivocation

Once the opposition between the two parties has been couched thus, it may still be claimed that which property is picked out by the two theories (or families thereof) is in its turn (partly) determined by the very principles of the theory or theories characterizing it. One may take such principles as what are often called "implicit definitions". At least in this sense, it may be claimed, the Meinongian and Quinean talk of different properties. When they speak with the vulgar, they mean the same by "exists". But when they engage in theoretical reflection on existence, they end up characterizing different properties. This weaker version of equivocation may be called "theoretical equivocation".

I think that theoretical equivocation, just as its stronger cousin, is mistaken. The analogy with debates on alternative logics may help. Classical logic and various non-classical logics oppose each other in their accounts of, say, negation: virtually any inferential feature of negation, from Double Negation Elimination to De Morgan's Laws and Minimal Contraposition, has been disputed by this or that logical party. So one sometimes hears the following view, reported, but not endorsed, by Graham Priest:

> There is no such thing as negation; there are lots of different negations: Boolean negation, intuitionist negation, De Morgan negation. Each of these behaves according to a set of rules (proof-theoretic or semantic); each is perfectly legitimate; and we are free to use whichever notion we wish, as long as we are clear about what we are doing. If this is right, there is nothing left to say about the question, except what justifies us in categorizing a connective as in the negation family. (Priest [21]: 76)

Priest thinks this view mistaken, and I think that theoretical equivocation on existence is mistaken for similar reasons. These may depend on some confusion

---

[13]See also the arguments in Evans [5], Ch. 10. Also for Evans "there seems to be very strong evidence that the English word 'exists' is used, at least on some occasions, to signify a first-level concept, true of everything" (345).

between a theory and what the theory is a theory of.[14] It would be wrong to claim that each of the two competing theories of existence (or families thereof), the Meinongian and the Quinean, characterizes its own object, let us say, *de dicto*, as whatever property satisfies the principles of the theory. Such a view makes a nonsense of debates in the foundations of logic – or, respectively, of ontology. It clashes with the belief, shared by the opposed parties, that they are disagreeing on substantial matters of fact. Meinongians and Quineans with diverging theories of existence manifest a flatly *de re* attitude. We have a Quinean characterization (or a family of characterizations) of a certain property, and a Meinongian one (ditto). The disagreement does not boil down to each party's arrogating the word "existence" to name its own property. Each apparently aims at theorizing on that property which is existence, providing a certain characterization of it. And each usually claims or implies that the other party's theorizing about the very same property is wrong. They have opposite theories, characterizing in incompatible ways the property of existence. For the Meinongian, *it* is a real property: a non-trivial feature that some things have, others lack. According to the (property-friendly) Quinean, *it* is not a real property: it is a trivial property anything has. Each party challenges the truth of the proposition expressed by (Q) and (M) respectively in the other party's mouth. "The theoretical object has to fit the real object; and how this behaves is not a matter of choice" (Priest [21]: Ibid). The party that happens to be wrong is wrong about the property that predicate picks out, both in its own mouth and in anyone else's; its being wrong is its having theoretical beliefs about that property, which are false. Again, we can keep claiming that the Quinean and the Meinongian have "different conceptions of quantification", and thus we can keep talking of "Meinongian quantification" as something distinct from "Quinean (standard, mainstream) quantification". But the sense in which Meinongians and Quineans oppose each other on quantification is the following: *of* the property of existence, the Quinean asserts that it is reducible to the commonly understood absolutely unrestricted (or, etc.) quantifier, whereas the Meinongian denies this.

11.5 Analyticity

Now that we have a plausible set-up for the debate, let us move on to the objection from analyticity. Unlike the objection from equivocation, this grants that (Q) and (M) express contradictory propositions in the mouths of the two parties. They

[14]Priest ascribes such a confusion to Quine, when in *Philosophical Logic* he makes his famous point on someone's disputing *ex falso quodlibet* for negation as "changing the subject": see Ibid, fn. 4. Williamson claims: "Quine's epistemological holism in Two Dogmas undermines his notorious later claim about the deviant logician's predicament" (Williamson [41]: 97).

encapsulate two theories of existence opposed on substantive matters of fact. However, the Meinongian's is an obviously flawed theory of existence. According to the aforementioned committed Quinean,

> Meinong's theory has a rather [...] important defect [...], and that is that it is self-contradictory – *obviously* self-contradictory. (Van Inwagen [37]: 39)

(M) is, in fact, only one small step away from logical self-refutation. Following Boghossian, call a sentence *Frege-analytic*, or such that it expresses a Frege-analytic truth, when it is synonymous with a logical truth, so that it can be obtained from the latter by replacement of synonyms (see Boghossian [2], referring to Frege [6], §3). "A bachelor is an unmarried man" is Frege-analytic: it can be obtained from the logical truth "A bachelor is a bachelor" by replacing the second occurrence of "bachelor" with the synonym "unmarried man". Then let us say that a sentence synonymous with a logical falsity, so that it can be obtained from the latter by replacement of synonyms, expresses a Frege-analytic falsity.

Now (M), the objection from analyticity goes, is one such sentence: "x exists" just is synonymous with "There is such a thing as x". So "It is not the case that x exists" is synonymous with "It is not the case that there is such a thing as x". So by one replacement of synonyms (M) turns into a logical inconsistency. Meinong himself specified the logical inconsistency at issue, by (in)famously claiming in his *Gegenstandstheorie*:

> Those who like paradoxical modes of expression could very well say: "There are objects of which it is true that there are no such objects." (Meinong [18]: 83)

This certainly looks like a *contradictio in terminis*. In the canonical notation, "...there is no such object (say, x)" becomes $\sim\exists y(y = x)$. Then, "There is some object, x, such that there is no such object x" becomes $\exists x\sim\exists y(y = x)$. "But the result of prefixing a tilde to this formula is a theorem of logic" (van Inwagen [33]: 38). Thus, Meinong's statement is a blatant logical absurdity. Compare how Frege himself makes the checkmate move in his dialogue with Pünjer on existence:

> [From your view, it] follows that there are objects of ideas – ideas which have not been caused by something affecting the ego – which do not exist. Now if you are using the word "exists" in the same sense as the expression "there is", then you have at the same time both asserted and denied the same predicate of the same subject. (Frege [7]: 65)

(Frege weakens the key claim into the antecedent of a conditional; but given his views on existence, the weakening looks like a rhetorical move: Pünjer is giving the same meaning to the two expressions, for he is a competent speaker; he just has a Frege-analytically mistaken view of existence).

In the current dialectical context, one cannot just *declare* that "x exists" is a synonym of "There is such a thing as x". One must provide evidence. Otherwise, the objection begs the question against the Meinongian. If existence is adequately captured by the quantifier, a logical notion, then to quantify on things of which it is said that they do not exist is a logical absurdity. But that to quantify is to commit to the existence of what one quantifies over, as we have seen, is denied by supporters of (M). For them, to exist is not to be the value of a variable. The Quineans raising the objection from analyticity typically point at some linguistic evidence, which they take as very stringent. The move may be named the Argument from Italics. It goes thus:

> Meinongianism entails that there are things [. . .] that have no being of any sort; but if there are such things, they obviously have being. For a thing to have *being* is for there to *be* such a thing as it; what else could being be? (Van Inwagen [37]: 39)

To claim that there (italics flagged:) *is* something that does not exist is to stumble upon the meaning of "there (flagged again:) *is*": no relevant difference can be detected between "there *is*" and "exists". What could "there is" stand for, as used in the vernacular, if not what "exists" stands for?

The objection from analyticity may have promoted a certain schism between Meinongians. In fact, two claims are made here to reach the conclusion that condemns (M):

(a) To quantify is to ascribe being to what one quantifies over.
(b) Being is the same as existence.

Taken together, (a) and (b) entail that to quantify is to ascribe existence to what one quantifies over. That "being is the same as existence" (van Inwagen [34]: 15) is another key claim of Quinean meta-ontology; but it is (a) that is supported by the Argument from Italics. Now some neo-Meinongians have felt the pressure of the Argument and, in order to block the conclusion that (M) is Frege-analytically false, they have denied (b): they have advocated a distinction between being and existence. They have accepted (a), granting that the expression "there is" brings with it commitment to the being of the things one quantifies over. But they have retained (M) by saying that these things have being in some form or other, despite lacking existence. Sometimes, this has been phrased as the claim that things lacking the full-fledged form of being we ordinarily call existence must have an impoverished or watered-down form of being, expressed by the "is" in "There is such a thing as *x*" (on this point, see e.g. Zalta [43]: 103–4).

This intermediate position seems to some extent a retreat: by attaching to quantification (watered-down) ontological commitment, commitment to the (watered-down) being of what one quantifies over, it looks close enough to (a watered-down) Quineanism. Matti Eklund calls this position "modes-of-being Meinongianism", and notices that Meinongianism is understood in this way "typically when the view is discussed by its foes" (Eklund [4]: 328), van Inwagen [37] being a good example. The Meinongian had better attack (a). She may accept that "being is the same as existence", and claim that some things just lack being i.e. existence (and if there are different ways or modes of being i.e. existence, they have none of them).[15] She may, thus, resist the Argument from Italics.

[15]"More often Meinongians instead hold the view that quantification is not ontologically committing in any sense. When I say that there are things that don't exist, among them Hamlet, I do not mean to ascribe being to Hamlet in any way" (Eklund [2006]: Ibid.). Eklund thus calls these "non-commitment Meinongians"; Priest [20] certainly is one of them.

How? I think the Meinongian will have to point at linguistic counter-evidence in order to oppose the alleged evidence highlighted by the Quinean. Considerations from ordinary language semantics, or meaning analysis, are not as popular today in analytic philosophy, and especially in metaphysics, as they were several decades ago. Few ontologists subscribe to Strawsonian "descriptive metaphysics", and some make of this a self-conscious methodological point.[16] But given the current dialectics, resorting to linguistic considerations on the Meinongian side seems a legitimate move. For it is the Quinean Argument from Italics that resorts to language now, in order to prove (M) Frege-analytically wrong. Again, we may look at analogous discussions on non-classical logics, where variations on the Argument from Italics can often be found. Tappenden [31] and Varzi [38] detect the widespread use of such arguments in debates on supervaluationism and its non-truth-functional account of disjunction: You claim that 'Either A or B' holds, so *either A or B* [stamp the foot, bang the table] must hold!". Or, here's one reaction to the dialetheic paraconsistent logician claiming that contradictory truth-bearers can both be true:

> [The dialetheist's] 'truth' is meant to be truth and his 'falsity' is meant to be *falsity*. More to the point, as above, his 'contradictories' are meant to be *contradictories*. Yet they cannot be, as we have seen. [. . .] While 'truth' and 'falsity' are only subcontraries in [the dialetheist's] language, that does not show, in any way, that truth and falsity are only subcontraries. For no change of language can alter the fact, only the mode of expression of them, as we saw before. And one central fact is that *contradictories* cannot be true together – by definition. (Slater [29]: 452–3)[17]

It seems clear that the alleged evidence pointed at by one who uses some version or other of the Argument from Italics against heterodox philosophical positions (Meinongianism, dialetheism, non-truth-functionalism, or else), is linguistic. What can the aim of stressing a piece of language by italicizing it be, in such contexts, if not to call the deviant theorist's attention on, well, its being there: "An occurrence of the verb 'to be' [stamp the foot, bang the table] is in (M). Can't you see it?". The reason why, granted that being is the same as existence, (M) expresses a Frege-analytic falsity and is "obviously self-contradictory", or only one substitution step away from logical falsity, would be the linguistic one that, well, there is an "is" in "there is".

---

[16] Todays ontologists are not conceptual analysts: few attend to ordinary usage of sentences like chairs exist. [. . .] Their methodology is rather quasi-scientific. They treat competing positions as tentative hypotheses about the world, and assess them with a loose battery of criteria for theory choice. Match with ordinary language and belief sometimes plays a role in this assessment, but typically not a dominant one (Sider [28]: 385).

[17] I admit I'm not really sure whether Slater has in mind and objection from equivocation here (which may be testified by his talk of "change of language"), rather than an objection to the effect that "Contradictories can be both true" is Frege-analytically false. His use of italics was too nice not to quote him anyway.

The reply to which is, tersely: that the verb "to be" shows up in some of the quantificational expressions we use lends thin linguistic support to the thick meta-ontological claim that we are always committed to the being-or-existence to whatever we quantify on.

11.6 Quantification and the Verb "to be"

Less tersely. It is commonly acknowledged that the verb "to be" of English (or its counterparts in other languages) can play quite different roles, and that this has been a source of puzzles for Western metaphysics since its origins (Bertrand Russell once stated that the ambiguity of "is" was a disgrace for the human race). One uses it, as an auxiliary verb, to form passives or to express progressive aspects; as one of the copulative verbs, to predicate something of something; or to express identity, inclusion or subsumption. Finally, we have a use of the verb that expresses existence. It is less widespread in English than it was for the founding fathers of Western metaphysics, the Greek philosophers.[18] It does show up though, often in philosophical, theological, or somewhat deep claims ("I think, therefore I am"; "God is"). The verb is used here, as is sometimes said, in an absolute way. Aristotle marked the distinction with the copulative use via the adverb $\alpha\pi\lambda\omega\varsigma$, corresponding to the Latin *simpliciter*, and which might be rendered as "being *tout-court*", or "being without qualification", in contrast with "being something" (being something-or-other, the having of properties expressed by predication, which a Meinongian would call *Sosein*):

> For it is not the same thing not to be something and not to be simpliciter, though owing to the similarity of language to be something appears to differ only a little from to be, and not to be something from not to be. (*On Sophistical Refutations*, 167a 4–6)

So "to be" can be used either absolutely (simpliciter, $\alpha\pi\lambda\omega\varsigma$), or not. When not, as when we speak of something's being identical with something, or of something's being affected by something, or in general of something's being such-and-such and so-and-so (*Sosein*), for the Meinongian this does not entail an ascription of being $\alpha\pi\lambda\omega\varsigma$, that is, existence, to these things. Holmes is identical with something – with Holmes. Holmes is something, he is such-and-such and so-and-so, i.e., he has

[18]As the Greeks did not have anything like our "exists" (a later Latin coinage) and different from (their counterpart of) "is", i.e., ™st. . . , they could only use the latter to express existence. In his *The Verb Be in Ancient Greek*, Charles Kahn lists absolute uses of *eînai* to express not only actual existence, but possibility (not only of the alethic, but also of the deontic kind: Kahn calls this the "potential construction"), and truth ("the veridical construction"): see Kahn [11]: 294, 336. Accordingly, "is not" was used not only to express nonexistence, but also impossibility or prohibition ("Is not to fight Zeus, son of Kronos", *Iliad* 21.193), and falsity ("This word of yours could [not] be", *Iliad* 24.56). Aristotle listed the veridical construction as expressing one way of being in the *Metaphysics*, but deferred its treatment to his works on logic, where it more properly belonged.

properties (the Sosein). He is: a character of *The Hound of the Baskervilles*, more famous than any existing detective, thought of by me, etc. Holmes, however, does not have being $\alpha\pi\lambda\omega\varsigma$ or existence, that is: if we check his various properties, the (non-blanket) one of being-or-existing turns out to be missing. When "to be" is used $\alpha\pi\lambda\omega\varsigma$, the Meinongian may grant that "being is the same as existence". What about that "is" in "there is", then? Isn't that "is" a case of "to be" being used $\alpha\pi\lambda\omega\varsigma$? The Meinongian may tell a story as to why this need not be the case. That there is an "is" in "there is" should not lead us to conclude that, in all cases, we ascribe being $\alpha\pi\lambda\omega\varsigma$ to the things we quantify over.

To begin with, quantificational devices have very different forms in natural languages; in lots of them, the verb "to be" or its counterparts in other languages don't show up at all. English uses "some" or "for some", where the verb "to be" does not appear, for the same purposes as "there is". The German often uses "*es gibt*" to express what we express in English by "there is"; but we would hardly conclude that, then, the Germans ascribe giving, or being given, to anything they quantify on, whatever "*given*" means in the context (in a sense, "*being given*" may anyway be closer than "exists" to expressing what one usually does by quantifying, as we are about to see). In French one typically says "il y a" for "there is". Here the French for "to be", "être", again does not show up. French mostly uses the verb "to have", "avoir", not the verb "to be", "être". Again, we wouldn't claim that the Frenchmen are ascribing having, avoir $\alpha\pi\lambda\omega\varsigma$ (?) to things, just because they quantify over them by using the expression "il y a". All kinds of verbs besides "to be", from "to give" to "to have", show up in quantificational constructions. In this sense, the verb "to be" is accidental to quantification.

Next such expressions, "there is", "for some", "il y a", "es gibt", "c'è", "si dà", etc., are often used as locative constructions. Their task is to present the relevant objects, to introduce them in discourse, or to situate them in a wider context. Because of this, they are often accompanied by explicit or implicit locational restrictions: "There are two trucks here, there are two more down there"; "There's a girl waiting in the car"; "There was a guy at the door this morning, looking for you". Locational restrictions are very often unfeasible with "exists". One wouldn't claim, of the four trucks, that two exist here while two exist down there; or that a girl exists in the car, waiting; or that a guy existed at the door this morning and, while he existed there, he was looking for you.[19] This is not a good result for a supposedly always-substitutable synonym.[20]

[19]As pointed out in Moltmann (2009, The Semantics of Existence, Unpublished MS), locational restrictions are acceptable in the case of mass nouns or bare plurals – things work much better here: "Lions exist both in Africa and in Asia"; "With such massive exploitation, soon oil will no longer exist in the Northern Sea".

[20] *Exist* applies to a subclass of entities that can be in the domain of *there*-sentences, excluding past and metaphysically possible (but not actual) objects, events, as well as intentional objects [. . .]. The general function of *there*-sentences appears to be to locate entities within either a larger domain of beings or a domain that is explicitly or implicitly restricted,

The Meinongian need not even rule out that "there is", in many if not most cases, *does* encode or entail being $\alpha\pi\lambda\omega\varsigma$, that is, existence. Existentially committing quantification is, for her, restricted quantification. The restriction can be specified via the appropriate (non-blanket) existence predicate, but can in most cases be left implicit, for it is conversationally understood. Contextual restrictions of quantification are common ground between Quineans and Meinongians. It is at times claimed that the Meinongian perspective has two quantifiers, the existentially neutral and the existentially loaded one, and that the Meinongian way is "the way of the two quantifiers" (van Inwagen [36]: 138). But this is a bit misleading: the existentially loaded quantifier just is the quantifier – restricted, explicitly or not, to existents. Also in this sense, "Meinongian quantification" may just be called quantification.

Contextual restrictions of quantification play various explanatory roles in ontology. Many mereologists, for instance, place no restriction on composition: given any x and y, we automatically have a z whose parts are exactly the parts of x and the parts of y. Hence we have such bizarre things as DKL's coins-Eiffel tower fusions mentioned above. The mereologist can explain our perplexity in front of such scattered, bizarre objects by saying that, typically, we don't quantify over them, for they are cognitively and practically irrelevant. In everyday talk, we tend to restrict our quantifiers to spatiotemporally handy things useful for our purposes. Those are the things that often count for us – we count on them – we quantify over them. We often focus on what Austin famously called "moderate-sized specimens of dry goods", disregarding very small, very large, discontinuous, ephemeral, or scattered things.[21]

Now an analogous remark holds for nonexistents. According to Meinong "The totality of what exists, including what has existed and will exist, is infinitely small in comparison with the totality of the Objects of knowledge" (Meinong [18]: 79). This smaller totality, however, is the set of things that matter more in our lives. He explained the "prejudice in favour of the actual" (and he meant: of the existent) as based on our "lively interest in reality which is part of our nature" (Ibid). It is no surprise that "there is" mostly brings commitment to the existence of what we quantify over. Indeed, we usually contextually refer only to the things that exist now and around here, or rather, only to some of them (i.e., our quantifiers are, usually, existentially loaded, presentist, restricted to our whereabouts and in various other

spatially, temporally, or otherwise, a function that seems to be reflected in the appearance of the locative *there*. (Ibid., Sect. 2.3).

[21] We are happy enough with mereological sums of things that contrast with their surroundings more than they do with one another; and that are adjacent, stick together, and act jointly. [. . .] We have no name for the mereological sum of the right half of my left shoe plus the Moon plus the sum of all Her Majestys ear-rings, except for the long and clumsy name I just gave it [. . .]. It is very sensible to ignore such a thing in our everyday thought and language. But ignoring it wont make it go away. (Lewis [13]: 211–13)

ways). Such default restrictions, though, can be abandoned: we say that some orcs from *The Lord of the Rings* look sinister, or that some of the cities Borges described are purely imaginary.[22]

11.7 Charitable Interpretation

I conclude with some general remarks on the two objections to Meinongianism. The (neo-)Meinongian theories try to show that the denial of Quineanism, that is, the claim that existence is not a blanket feature of everything, is ontologically, epistemically, and semantically coherent. Such theories, as is true of most theories, may in the end fail. But attempting to refute them by appealing to a presumed Frege-analyticity of (Q), or to the fact that the mere understanding of the proposition expressed by it is sufficient to mandate assent, is, I suspect, not a very promising move. Many ontological views can be easily accused of equivocation or analytical falsehood. In the ontology of mathematics, one could reject Platonism in one move: it is analytic that everything is concrete. In modal metaphysics, one could similarly reject DKL's modal realism: it is analytic that everything is actual; so Lewis' ontology of nonactual *possibilia* is analytically false. In the early days of modal realism, some have tried such a fast rejection.[23] But modal realism and mathematical Platonism nowadays are respectable theories on the philosophical market. Many philosophers believe them to be false, and even necessarily so, given the modal status of their key claims; but to show this, one needs more than a quick argument to the effect that such key claims are analytic falsehoods, or denials of things that command assent to competent speakers of English. It seems to me that the issues addressed in this paper, too, show how alleged knock-down arguments imputing obvious mistakes rarely achieve their expected results. Conversely, the most promising attempts at refuting a philosophical view may often start by assigning to their target as much theoretical strength as charitable interpretation allows.

[22] Along the same lines, see also Salmon [27]: 56–7. The Medieval logicians conception of quantification smoothly dealt with such contextual domain expansions (see e.g. Ashworth [1], Read [24], Priest [20], Section 3.7). In the mainstream doctrine of the *suppositio terminorum*, "Some S is P" is by default true if and only if something that is actually currently S is P. But the ordinary *suppositio* can be expanded in intensional, modal, temporal contexts to possible, past, future objects not currently or actually existing: "Some S has been P", is true if and only if something that is or has been S, is or has been P, even though it does not exist now ("Some monks have been knights"). "Some S can be P" is true if and only if something that is or could be S is or could be P, even though it is merely possible ("A golden mountain can be as large as Mount Ventoux").

[23] For example, Richards [25], Haack [9], and Lycan [15].

Acknowledgements Parts of this paper were presented at the *2013 Tübingen Metaphysics Workshop*: thanks to the organizers Thomas Sattig and Alessandro Torza for having me, and to Graham Priest, Jason Turner, Boris Hennig, Bjørn Jespersen, Friederike Moltmann, and especially to Tuomas Tahko as the discussant of my talk, for their useful comments and remarks.

References

1. Ashworth, E.J. 1977. Chimeras and imaginary objects: A study in post-medieval theory of signification. *Vivarium* 15: 57–79.
2. Boghossian, P. 1997. Analyticity. In *A companion to the philosophy of language*, ed. B. Hale and C. Wright. Oxford: Blackwell.
3. Burgess, J., and Rosen, G. 1997. *A subject with no object*. Oxford: Oxford University Press.
4. Eklund, M. 2006. Metaontology. *Philosophy Compass* 1: 17–34.
5. Evans, G. 1982. *The varieties of reference*, ed. J. McDowell. Oxford: Oxford University Press.
6. Frege, G. 1884a. Die Grundlagen der Arithmetik. Eine logisch-mathematische Untersuchung über den Begriff der Zahl. Hamburg: Felix Meiner Verlag, 1986.
7. Frege, G. 1884b. Dialog mit Pünjer über Existenz. In ed. Kleine Schriften, Darmstat, Hildesheim, 1967: 60–75. Tr.: Pünjer on Existence. In ed. Hermes, Kambartel and Kaulbach, 1979. Frege: Postumous Writings, Blackwell, Oxford: 53–67.
8. Grossmann, R. 1974. Meinong. *The arguments of philosophers*. London/New York: Routledge.
9. Haack, S. 1977. Lewis’ ontological Slum. *Review of Metaphysics* 33: 415–429.
10. Hirsch, E. 2002. Quantifier variance and realism. *Philosophical Issues* 12: 51–73.
11. Kahn, C.H. 1973. *The verb “be” in ancient Greek*. Dordrecht: Kluwer.
12. Kant, I. 1781. Kritik der Reinen Vernunft. In *Gesammelte Schriften*, ed. der Königlich Preußischen Akademie der Wissenschaften, de Gruyter & Co., Berlin 1969. Tr.: Critique of Pure Reason, The Humanities Press, New York 1950.
13. Lewis, D.K. 1986. *On the plurality of worlds*. Oxford: Blackwell.
14. Lewis, D.K. 1990. Noneism or allism? *Mind* 99: 23–31.
15. Lycan, W. 1979. The trouble with possible worlds. In *The possible and the actual*, ed. M.J. Loux, 274–316. Ithaca: Cornell University Press.
16. Marconi, D. 1997. *Lexical competence*. Cambridge: MIT.
17. McGinn, C. 2000. Logical properties. In *Identity, existence, predication, necessity, truth*. Oxford: Oxford University Press.
18. Meinong, A. 1904. Über Gegenstandstheorie. In *Untersuchungen zur Gegenstandstheorie und Psychologie*, ed. A. Meinong, 1–51. Leipzig: J.A.; Barth. Tr.: 1960. The theory of objects. In *Realism and the background of phenomenology*, ed. Chisholm, 76–117. New York/London: The Free press-Collier-Macmillan.
19. Priest, G. 1995. *Beyond the limits of thought*. Cambridge: Cambridge University Press. 2nd, exp. ed. Oxford University Press, Oxford 2006.
20. Priest, G. 2005. *Towards non-being. The logic and metaphysics of intentionality*. Oxford: Oxford University Press.
21. Priest, G. 2006. *Doubt truth to be a liar*. Oxford: Oxford University Press.
22. Quine, W.V.O. 1948. On what there is. *Review of Metaphysics* 2: 21–38. Reprint in: From a logical point of view. Cambridge: Harvard University Press.
23. Rayo, A., and G. Uzquiano. (eds.) 2006. *Absolute generality*. Oxford: Oxford University Press.
24. Read, S. 2001. *Medieval theories of properties and terms*. The Stanford encyclopedia of philosophy. Stanford: CSLI. http://plato.stanford.edu/entries/medieval-terms.
25. Richards, T. 1975. The worlds of David Lewis. *Australasian Journal of Philosophy* 53: 105–118.
26. Routley, R. 1980. *Exploring Meinong’s jungle and beyond*. Canberra: Australian National University RSSS.

27. Salmon, N. 1987. Existence. *Philosophical Perspectives* 1: 49–108.
28. Sider, T. 2009. Ontological realism. In *Metametaphysics*, ed. D. Chalmers, D. Manley, and R. Wasserman (2007), 384–422. Oxford: Oxford University Press.
29. Slater, B.H. 1995. Paraconsistent logics? *Journal of Philosophical Logic* 24: 451–454.
30. Stanley, J. 2001. Hermeneutic fictionalism. *Midwest Studies in Philosophy* 25: 36–71.
31. Tappenden, J. 1999. Negation, denial and language change in philosophical logic. In *What is negation*? ed. D.M.Gabbay and H. Wansing, 261–298. Dordrecht: Kluwer.
32. Thomasson, A. 2009. Answerable and unanswerable questions. In *Metametaphysics*, ed. Chalmers, Manley, and Wasserman (2007), 444–470. Oxford: Oxford University Press.
33. Van Inwagen P. 1977. Creatures of fiction. *American Philosophical Quarterly* 14: 299–308. Reprint in: van Inwagen. 2001. *Ontology, identity and modality. Essays in metaphysics*, 37–56. Cambridge: Cambridge University Press.
34. Van Inwagen P. 1998. Meta-ontology. *Erkenntnis* 48: 233–250. Reprint in: van Inwagen. 2001. *Ontology, identity and modality. Essays in metaphysics*, 13–31. Cambridge: Cambridge University Press.
35. Van Inwagen P. 2001. *Introduction to ontology, identity and modality. Essays in metaphysics*, 1–10. Cambridge: Cambridge University Press.
36. Van Inwagen P. 2003. Existence, ontological commitment, and fictional entities. In *Metaphysics*, ed. M. Loux and D. Zimmermann, 131–157. Oxford: Oxford University Press.
37. Van Inwagen P. 2006. McGinn on existence. *The Philosophical Quarterly* 58: 36–58. Reprint in: Bottani and Davies (eds.) 2006. *Modes of existence. Papers in ontology and philosophical logic*, 105–129. Frankfurt: Ontos.
38. Varzi, A.C. 2004. Conjunction and contradiction. In *The law of non-contradiction*, ed. J.C. Beall, G. Priest, and B. Armour-Garb, 93–110. Oxford: Clarendon.
39. Varzi, A.C. 2005. *Ontologia*. Rome: Laterza.
40. Williamson, T. 2003. Everything. *Philosophical Perspectives* 17: 415–465.
41. Williamson, T. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell
42. Wolstertorff, N. 1961. Referring and existing. *The Philosophical Quarterly* 11: 335–349.
43. Zalta, E. 1988. *Intensional logic and the metaphysics of intentionality*. Cambridge: MIT.

Chapter 12
Qualifying Quantifying-in

Bjørn Jespersen

Abstract Quantifying-in is existential quantification into non-extensional contexts headed by a modal or attitudinal operator. The sense and sensibility of quantifying-in has often been challenged. This paper outlines a transparency-preserving semantics as a prerequisite for the logical validity of quantifying-in. The paper demonstrates how to formally validate quantifying into a non-factive, hyperintensional attitude context.

The rule of existential generalization is one of the hallmarks of extensional logic. It is also a rule with strong intuitive appeal. Its conclusion makes explicit an ontological commitment implicit in the premise. If the premise states that some particular a has some specific quality F then the conclusion states that there is some x that has F. For instance, if the individual Tilman lives in Tilburg then there is at least one element x in the domain of individuals who lives in Tilburg. If the premise is true then the conclusion is the truth that the *quantity* of objects with the particular *quality* of living in Tilburg amounts to at least one. In set-theoretic terms, the set of individuals living in Tilburg is said to be non-empty. The conclusion is indifferent to whether it is Tilman or some other individual who lives in Tilburg: *any* individual will do as long as some individual has the quality in question. The idea underlying quantification, whether existential or universal or generalized (e.g. *at least three*), is abstraction from the particular and specific to the general and arbitrary in order to extract a pure quantity. It would be misconceived to ask which particular element of the domain of quantification is the value of the existentially bound variable.

B. Jespersen (✉)
LOGOS, Departament de lògica, història i filosofia de la ciència, Universitat de Barcelona,
C/ Montalegre 6, 08001, Barcelona, Spain
e-mail: bjorn.jespersen@gmail.com

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_12

So far, so good. This inference schema is an uncontroversial component of Logic 101:

$$(1)\ \frac{Fa}{\exists x(Fx)}\ \text{EG}$$

We may haggle over the formal semantics of the schema, e.g. whether to project an objectual versus substitutional, or a constructivist versus realist, etc. interpretation of the quantifier onto the schema. But whatever the details of the logic of the inference and the meaning of the conclusion, the very validity of the schema is beyond dispute. Its validity is a datum that any viable semantics of quantifiers must accommodate. We need stray only a bit from (1), however, before dispute erupts. For with "*Fa*" as our premise, there are two positions to quantify into—the position of the singular term 'a' and the position of the predicate '*F*'. If Tilman lives in Tilburg, as before, then he is not someone without properties, for there is at least one property f that Tilman has:

$$(2)\ \frac{Fa}{\exists f(fa)}\ \text{EG}$$

The philosophical objection is that quantifying over properties goes too far. Tilman may well live in Tilburg, but this is not to say that he thereby has a property, for properties form an ontological category that is better dismissed. This sort of objection tends to feed on a general distrust in intensional entities such as properties, propositions, magnitudes (e.g. *the number of planets*), individual roles (e.g. *the Sultan of Brunei*), etc.[1]

Existential generalization becomes even more contentious, as well as more complicated, than (2) when quantification mixes with modalities or attitudes. When this happens, one or more existential quantifiers need to bind one or more occurrences of one or more variables inside the scope of one or more modal or attitudinal operators (functors). Restricted to one schematic modal or attitudinal operator *Op* and one quantifier, the characteristic scope distribution is this[2]:

$\exists x \ldots Op \ldots Fx$

Here is a semi-formalized example:

$$(3)\ \frac{\text{Contingently, } Fa}{\exists x(\text{contingently, } Fx)}\ \text{EG}$$

[1]Provided both schemata are valid, it is an option to unify them into a third schema, whose conclusion means that somebody (something) has some property:

$$(1{+}2)\ \frac{Fa}{\exists x \exists f(fx)}\ \text{EG}$$

[2]The inverse scope distribution, $Op \ldots \exists x \ldots Fx$, would be the schema underlying sentences like "Contingently, there is some x such that x is an F". This scope distribution does not exemplify quantifying-in and, while being of great independent logical and philosophical interest, will not be discussed here.

(3) exemplifies so-called *quantifying-in*. This is so because the ∃-bound occurrence of x occurs inside the scope of the modal operator 'contingently'. Similarly for f in (4):

$$(4)\ \frac{\text{Contingently}, Fa}{\exists f(\text{contingently}, fa)}\ \text{EG}$$

The first at least well-known modern-day example of quantifying into modalities is presumably the Barcan Formula:

$$\Diamond\exists xFx \rightarrow \exists x \Diamond Fx$$

Where $\Diamond$ represents logical possibility, *BF* states that if it is logically possible that something be an F then something has the logical potential to be an F. The ∃-bound occurrence of x in the consequent falls within the scope of $\Diamond$, so the question arises how ∃ succeeds in reaching across $\Diamond$ and binding this occurrence. $\Diamond Fx_i$ is a modal, hence intension-sensitive context, so if ∃ binds x_i that range over extensional entities such as individuals, the question arises whether the occurrences of x_i within the context $\Diamond Fx_i$ retains their quantificational range or must range over individuals-in-intension or are maybe deprived of a range altogether.[3]

The earliest discussion of what is in effect quantification into *attitudes* may be one of Buridanus's *insolubilia* from 1350 which Geach [11, p. 430] represents thus[4]:

> Let us then have our horse-coper arguing again. "If I owe you a horse, then I owe you something. And if I owe you something, then there is something I owe you. And this can only be a thoroughbred of mine: you aren't going to say that in virtue of what I said there's something else I owe you. Very well, then: by your claim, there's one of my thoroughbreds I owe you. Please tell me which one it is."

Of course, there is no one particular horse the coper owes to the man he is trying to befuddle with a fallacy. The coper's trick is the transition from there being a particular property that the coper owes the man an arbitrary instance of to there being a particular instance of that property that the coper owes.

Quine famously challenged quantified modal logic to make quantifying-in comprehensible. He later went on to challenge attitude logic to do the same. The first modern example of mixing existential quantification with attitudes is probably Quine [20]. The topic of quantifying into attitudes arises for Quine due to his notion of *relational attitude*. If Quine wants a sloop then, if his attitude is notional, any sloop will do to relieve him of slooplessness, and if his attitude is relational then only a particular sloop will satisfy his wish. The latter is arguably ambiguous: Quine wants a particular object, which happens to be a sloop; or Quine has his mind set on one particular sloop, to the exclusion of all other sloops.[5] Be that as it may, Quine phrases relational attitudes by means of quantified locutions: "There is an x such

[3] See Williamson [32, pp. 46ff] on the early reactions to quantifying into *BF*.

[4] See Zimmermann [33, p. 715, n. 1].

[5] See Sainsbury [21].

that x is a sloop and Quine wants x". It is this phrasing of relational attitudes that, historically at least, pushes quantifying-in to the fore in attitude logic.

Quine is dismissive of quantifying into modal contexts, for fear of condoning modality de re. Quine is also dismissive of quantifying into attitude contexts exemplifying notional attitudes. He is, as we just saw, sympathetic to quantifying into contexts exemplifying relational attitudes.[6] Quine's problem with this particular category of attitudes is how to make good logical sense of it. This leads him to put forward his well-known three-place analysis, which is intended to preserve referential transparency. Whatever one makes of this analysis, at least part of the appeal of relational attitudes seems to be that there is guaranteed to be an individual at the receiving end of the attitude: there is an individual whom or which the attitude is directed at (hence the qualification 'relational').[7] There, at least, quantification over individuals seems safe.

Quantification over individuals also at least appears to be safe as soon as the modal or attitudinal context being quantified into is a *factive* one. Necessity and contingency are factive modalities: from "Necessarily, A is true" and "Contingently, A is true" we can infer "A is true". From "Possibly, A is true" we cannot infer "A is true", but only "Maybe A is true". Knowledge is a factive attitude: from "a knows that A is true" we can infer "A is true". Belief is not factive: from "a believes that A is true" we cannot even infer "Maybe A is true", for we have not excluded the possibility that A be inconsistent. From "Necessarily, Venus has a moon" we can infer "There is an individual x such that x is a moon of Venus". (Never mind soundness; validity is what we are after.) From "a knows that Venus has a moon" we may likewise infer "There is an individual x such that x is a moon of Venus". Some non-propositional attitudes are also factive. For instance, from "a finds the site of Troy" we may infer "There is an individual x such that x is the site of Troy". But from "a seeks the site of Troy" we may not infer this conclusion, for there may be no unique site of Troy.

In general, two issues bearing on quantifying-in need to be kept separate. One is quantification *over* the domain that b belongs to, where b occurs inside the context $Op \dots b \dots$. The other is quantification *into* the context $Op \dots c \dots$.. The former makes explicit the ontological commitment mentioned at the outset, specifying the ontological category of b. The latter sort of conclusion extracts a component c from the attitude complement, which is the scope of Op, and quantifies over its domain: $\exists z \dots Op \dots z \dots$.. One reason for operating with this distinction is that we may, for instance, quantify *into* a *hyperintensional* context and quantify *over* a particular kind of *intensional* or *extensional* entity. Another reason is that we may want to quantify

[6] See not least Kaplan [16, pp. 230–31] and Crawford [3].

[7] Of course, it is not just that there is an individual at whom or which the attitude is directed—there is *one particular* individual at the receiving end of the attitude. It is doubtful whether existential quantification possesses enough expressive power to capture the dimension of particularity, as in "There is a princess that the prince wishes to marry, though not just any princess out there, but one particular princess." Cf. my initial remark about the quantity of instances of a particular quality or property.

into one, but not all, positions inside a given context. For instance, if *a* is seeking an abominable snowman then there is a *z* such that *z* is abominable and *a* is seeking *z* (i.e. *a* is seeking somebody or something abominable). The salient question is here what the quantificational range of *z* may be. If only extensional entities like individuals are an option then there will be far fewer cases of quantifying-in. The number of cases shoots up if also *intensional* entities, i.e. functions from possible worlds, are declared legitimate quantificational ranges. The number further increases if also *hyperintensional* entities may be quantified over. I will show that when we are quantifying *into* a hyperintensional empirical attitude context, such as believing that the glass before you is half-empty (but not necessarily also that the glass is half-full, or vice versa), then we may always quantify *over* hyperintensions and intensions, but not individuals, due to the non-factivity of doxastic attitudes.[8] That is, when *a* is seeking an abominable snowman, or some other cryptozoid, no instance of the property of being an abominable snowman has been quantified into existence, so to speak.

In this paper I will address three topics. The first two belong together, one being a special case of the other. The first topic concerns quantifying into *non-factive* attitude contexts, as just sketched. The second topic concerns what I dub *doublethink*, borrowing a term from Orwell, to describe inconsistent beliefs, which are inherently non-factive. (I have no conceptual space for dialetheia.) If quantification is restricted to extensions, in neither case is quantifying-in warranted, which is only reasonable, since Quine seems to have fielded the (im-) possibility of quantifying-in as something like a nonsense-detector: does the application of EG to premises involving attitudes or modalities eventuate in nonsensical or otherwise untoward conclusions? Both of the above cases call for qualification to warrant quantifying-in. I will be describing both topics mainly in prose, referring to existing literature that contains the formal details of my approach. The third topic is technical in nature, exemplifying how to quantify into a non-factive, empirical, hyperpropositional attitude context and quantifying over intensional entities.

This paper is the latest installment of a series of papers devoted to quantifying-in. The other papers are Duží and Jespersen [5], Duží and Jespersen [6], and Duží and Jespersen [7]. These papers themselves build primarily on Tichý [30], Materna [18], and Duží et al. [8, §5.3]. The present paper offers more by way of philosophical exposition and critique than was possible in the three papers co-authored with Duží, which target mainly the logical and semantic intricacies of quantifying-in.

The rest of the paper is organized as follows. Section 12.1 clears the semantic ground for the logic of quantifying-in. Section 12.2 discusses how to qualify

[8]The need for hyperintensions as the complements of at least some attitudes was realized in modern times as early as Carnap [2, pp. 53–54], thus it was known from its very inception that modal logic was insufficient as an all-encompassing attitude logic. Marcus, in a 1961 paper, also notes that epistemic and doxastic attitudes require 'a stronger equivalence relation… than strict equivalence', mentioning Carnap's *intensional isomorphism* as one attempt to obtain what is in effect a hyperintensional criterion of equivalence (Marcus [17, p. 14]).

non-factive attitudes, including inconsistent beliefs, so as to warrant quantifying-in. Section 12.3 demonstrates the logical details of quantifying into a hyperintensional context, as described above.

12.1 Top-Down Transparency

It is important to bear in mind that Quine does not reject quantifying-in tout court. Rather he points out that quantifying-in is problematic. But he never unambiguously identifies what the source, as opposed to symptom, of the problem is. Is it failure of Leibniz's Law? The existence of opacity? The mix of quantification and modalities or attitudes within the same context? As I suggested above, most likely quantifying-in is fielded as a criterion or stress test for whether modal and attitudinal contexts behave logically, which for Quine means behaving according to the laws of extensional logic. If existential generalization fails for modal or attitudinal contexts then he has constructed an argument with a false or even nonsensical conclusion, hence there must be something illogical about either one or more of the premises, or one or more of the rules being invoked, and the culprits have no place in an austere, regimented language. What I call the argument from doublethink is modeled on Quine's prior modal arguments also designed to derive the absurd conclusion that some number x has, and also lacks, a modal property such as being necessarily larger than five, or that some individual y has, and also lacks, a modal property such as being necessarily two-legged.[9] A doublethink *reductio* is pivoted on quantifying into existence an individual that somebody believes, and also fails to believe, to be an F.

Here is a reconstruction of one way of going about generating such a Quine-style *reductio*. One deploys an extensionalist semantics to contexts that are sensitive to more than just the identity and difference of extensions, and Leibniz's Law, *as defined for extensional contexts*, turns out to be invalid (not surprisingly). One then concludes that non-extensionalist contexts are logically lawless, or at the very least iffy. Quine's general stance is that any sort of context that defies substitution of identicals (i.e. is opaque) must also defy quantifying-in.

Kaplan, in reaction to Quine, pursues a different stance. Kaplan [16, pp. 242ff] maintains that failure of intersubstitutivity does not entail failure of quantifying-in, deploying his theory of arc quotes to get this project off the ground.

Forbes [10] sees it as his task to conceive of a device that will bar co-referential names from being substitutable within certain contexts in order to create 'substitution-resistant positions' (ibid., p. 352) without altering their reference

[9] See Duží et al. [8, §4.2.1] and Marcus [17, pp. 18–21]. Quine would not phrase his point in terms of properties, but I do, in order to extract a general lesson that applies also beyond the Procrustean bed of Quine's extensionalist semantics.

relation.[10] Forbes (ibid., pp. 357–62) does sketch a template for 'unproblematic quantifying-in', which is predicated on exportation of, in this case, 'Superman'[11]:

(a) Lois believes that Superman is an extraterrestrial
(b) Superman is someone whom Lois believes to be an extraterrestrial

---

(c) There is someone whom Lois believes to be an extraterrestrial

But the transformation (as Forbes ibid., p. 358 calls it) of (a) into (b) obliterates the huge differences between de dicto and de re attitudes.[12] So I am not convinced that Forbes has shown how to, indirectly, quantify into (a).

Whatever the details, I believe all three approaches are conceptually misguided.[13] The way I see it, the rules of extensional logic, including existential generalization, referential transparency, substitutivity of identicals, and compositionality form one package, such that all of them must be accommodated simultaneously. Cherry-picking is not a viable option. In essence, what is wanted is an extensional logic of non-extensional (including hyperintensional) contexts.[14] A semantics heeding universal transparency must precede such a logic. The philosophical idea is that once universal transparency has been safeguarded we have made a critical step toward availing ourselves of an extensional logic for any and all sorts of contexts, i.e. the extensional, the intensional, and the hyperintensional ones. Any such logic will, qua extensional, validate the rules of extensional logic, including EG. We need to put behind us the idea that a logic for non-extensional contexts must itself be non-extensional.[15]

[10]Forbes's device is *logophors*, so-labeled. 'Superman' and 'Clark Kent' refer to the same individual, but when reporting Lois's attitude toward this individual it affects the truth-value of the report whether 'Superman' or 'Clark Kent' is used to report her attitude. Forbes's logophors appear to be more quotational in character than Forbes would want, being an uneasy halfway house between Quinian and Fregean tenets. See Forbes [9] for his most recent application of logophors.

[11]Furthermore, from (b) together with (d) Superman is Clark Kent, we are supposed to infer (e) Clark Kent is someone whom Lois believes to be an extraterrestrial.

[12]See Duží and Jespersen [5].

[13]So is the approach in Priest [19]. Priest argues that substitutivity of identicals fails for both modal and epistemic contexts, and therefore sees it as his task to construct a logic (based around impossible worlds) in which the rule does not hold in order to accommodate arguments such as this: (i) This man [wearing a hood] is your brother; (ii) You do not know who this man is; (iii) Therefore, you do not know who your brother is. In my view, Priest's analysis of (i) is too heavy on the extensions, leaving no conceptual space for convergence between intensions. (See also Footnote 17 below.)

[14]See Duží [4].

[15]Davidson's sketch of his so-called paratactic theory of attitude contexts is on the same track. Davidson likewise eschews reference shift in order to heed 'semantic innocence'. For a brief comparison between Davidson's 'paratactic' approach and the 'hypotactic' one I am advocating here, see Duží et al. [8, p. 12]. See also Bealer [1, p. 148].

Sylvan [25, p. 29] cites an original example of alleged opacity, which none the less replicates a familiar pattern:

> The denominator of $\frac{2}{4}$ is 4. But $\frac{2}{4} = \frac{1}{2}$. So by transparency [substitution of identicals], the denominator of $\frac{1}{2}$ is 4.

Which, of course, it is not. Sylvan uses this fallacy to argue that not all mathematical contexts are extensional and transparent, some being intensional and opaque. I think Sylvan overstates what goes wrong in the above substitution. He is right that the conclusion does not follow. But this is not due to opacity; it is due to *wrong substituends*. In this sense the fallacy is about non-extensionality, because extensional values cannot be swapped:

(i)	*Denominator* $(\frac{2}{4}) = 4$
(ii)	$\frac{2}{4} = \frac{1}{2}$
(iii)	*Denominator* $(\frac{1}{2}) = 4$

When the denominator of a fraction is highlighted, as per (i) and (iii), a sensitivity to computational or algebraic structure is highlighted that is absent when only the result of computing the fraction is relevant, as per (ii). So where exactly does the argument go wrong? As almost always, an identity or equivalence at a lower level, (ii), is transferred up into a higher-level context, (iii). To be sure, two quarter dollars is the same amount as a half-dollar, but these are two different ways of arriving at the same amount of fifty cents. (i) and (iii) are sensitive to the differences between these two different ways of arriving at fifty cents, whereas (ii) is not. Hence inside the context induced by 'the denominator of' the only sort of appropriate substituends are fractions with the same denominator as in the original premise, (i). This makes for an exceedingly exacting criterion of substitutivity. The proper conclusion requires that the denominator be 4, while leaving room for a numerator different from 2. Whether a different numerator is possible *salva veritate* is a mathematical matter and not a logical one (bearing on preservation of validity) or a semantic one (bearing on preservation of meaning).

The well-known pattern Sylvan's example replicates is this. Already Smullyan [23] objected to this example of Quine's (adjusted to our post-Plutonic times), stressing the need to distinguish between a condition and the satisfier of the condition:

(i.i)	Necessarily, 8 exceeds 5
(ii.i)	The number of planets is 8
(iii.i)	Necessarily, the number of planets exceeds 5

The argument is valid, *provided* 'The number of planets' names 8, just as '8' does. If this is so then (ii.i) states the self-identity of a number co-denoted by a definite description and a constant. So (iii.i) and (i.i) come out denoting one and the same proposition, though phrasing it differently, and the conclusion is identical to one of the premises. If the argument is going to pack any punch, (ii.i) must receive a somewhat different analysis. It is most reasonably construed as stating a logically *contingent* astronomical fact, which may still be grounded in nomological necessity:

as the laws of nature would have it, the number of planets is 8, though logically it might have been another number, including 0. Accordingly, the number of planets is a *magnitude* (an empirical condition) whose values (satisfiers) are natural numbers. (ii.i) means that, contingently, the magnitude denoted by 'The number of planets' takes the value 8. Therefore, 'The number of planets' cannot replace '8' inside the scope of '□', which requires equivalence of conditions and not just of their satisfiers.[16]

Another example is *Partee's puzzle*:

(i.ii)	The temperature is rising
(ii.ii)	The temperature is 90°F
(iii.ii)	90°F is rising

The first premise ascribes the property of rising to a magnitude (the temperature at some location), whereas the second premise picks out the value of the magnitude (at a particular index that is suppressed). The context '...is rising' requires as a substituend a term for a magnitude rather than just a term for one of its values.[17]

Kaplan's example below buttresses my suspicion that an insufficient analysis of (what appears to be) identity sentences fuels much of the frustration with operating on modal and attitude contexts. Kaplan [16, p. 264] considers this argument:

(i.iii)	It will soon be the case that the President of the United States is a woman
(ii.iii)	The President of the United States = Nancy Reagan's spouse
(iii.iii)	It will soon be the case that Nancy Reagan's spouse is a woman

The conclusion is open to two readings, at least, on both of which it is rather peculiar, as intended by Kaplan. Either 'Nancy Reagan's spouse' is taken to have as its semantic value the individual who (in the mid-1980s) is both the President of the United States and the husband of Nancy Reagan, or whatever individual is a woman and Nancy Reagan's spouse in the near future (counting from 1986). The first reading we may call the transsexual one, for it requires that Nancy Reagan's spouse become a woman. The situation on the ground is that Ronald Reagan, as it

[16]This argument is valid, because we remain safely within the same sort of context throughout (and the mathematics checks out): □(8 exceeds 5); □($2^3 = 8$) / □ (2^3 exceeds 5). See also Marcus [17, pp. 36–38] on Smullyan. It is a by now well-established insight that "the equality relation that holds between expressions such as '9' and 'the number of planets' must be distinguished from the equality relation that holds, for example, between the expressions '9' and '7 + 2'. (Marcus, ibid., p. 37.) (I would prefer the equality relations to relate, in the final analysis, not expressions, but their non-linguistic denotations.) In particular, this expresses a falsehood: "□ (9 = the number of planets)", even if pretending that our solar system has exactly 9 planets, for this necessitation is not an option, as soon as "9 = the number of planets" is taken to state a contingent astronomical fact.

[17]For further details and discussion, see Duží et al. [8, pp. 124–25].

happens, will be undergoing radical surgery. The second reading is the lesbian one, for it requires Nancy Reagan to go ahead and marry a woman. Kaplan's verdict is:

> Thus, substitutivity fails. Contexts of *S* [the operator 'it will soon be the case'. Author] are opaque. (Ibid.)

I agree that substitutivity fails—for the *wrong* sort of substituends, that is. And if opacity is immunity to substitutivity of wrong substituends then opacity is a good thing. What is going on, though, is that Kaplan misdiagnoses the fallacy of the argument. The problem is that (i.iii) and (iii.iii) are explicitly temporal while (ii.iii) is not. The proper temporalization of (ii.iii) would be along the lines of:

(ii.iii') Currently, the President of the United States = Nancy Reagan's spouse

If we set the current time as t_0, and the near future at t_1, then 'The President of the United States at t_0' and 'Nancy Reagan's spouse at t_0' share the same semantic value; Ronald Reagan, given the actual course of events. On the transsexual reading, (iii.iii) goes into "At t_1, the President of the United States at t_0 is a woman". On the lesbian reading, (iii.iii) goes into "At t_1, Nancy Reagan's spouse at t_1 is a woman".

But we also need to analyze '=' properly.[18] It obviously does not state the self-identity of some individual who is both the President of the United States and Nancy Reagan's spouse. So what does it state? "The President of the United States at t_0 = Nancy Reagan's spouse at t_0" comes with temporalization and is on the right track. But the analysis is an analysis of time-indexed definite descriptions: given a time, a definite description denotes an individual (or nothing at all), on a Kaplan-style extensionalist analysis of definite descriptions. This analysis leaves it obscure what the semantics is of a definite description in the absence of a time assignment. In a paper published the same year as Kaplan's, Tichý [30, p. 254] says,

> The sentence ["The man who lives next door is the man who runs the city"] conveys information about two offices, that of the man who lives next door and that of the man who runs the city. It gives us no clue as to who occupies those offices. But it tells us nevertheless something about them that might not have been the case: namely that they are co-occupied, that some individual or other holds them both. We have seen [ibid., p. 253] that an office is a function whose value at a world-time is the occupant (if any) of the office in that world at that time. The assertive content of an identity sentence like the one just considered is simply to the effect that two such functions happen to take the same value in the actual world at the present time.

For starters, then, these two identity statements must be kept separate:

(i) *office*$_1 =_i$ *office*$_2$ *(identity between intensions)*
(e) *office*$_1$*(wt)* $=_e$ *office*$_2$*(wt) (identity between extensions)*

[18] Sleigh [22] takes opacity to be a fact of linguistic life, noting the resulting standard issues with substitutivity of identicals and quantifying-in. However, Sleigh offers no analysis of "Cicero = Tully" (ibid., p. 23) or "Dr. Salazar = the dictator of Portugal" (ibid., p. 24), nor does he note their obvious differences.

In Tichý's semantic theory the analysis of (ii.iii) must define the set of world-time pairs at which the two offices named 'The President of the United States' and 'Nancy Reagan's spouse', respectively, share the same occupant. The analysis does not include Ronald Reagan. The set of world-time pairs at which those two offices are co-occupied will have as a proper subset those world-time pairs at which Ronald Reagan is their shared occupant.

Tichýs' semantics is designed with universal transparency in mind.[19] As far as definite descriptions are concerned, his basic tenet is that they nowhere and never denote their respective unique descriptum (if any) and instead always, in any sort of context, denote one and the same empirical condition (what he calls an 'office'), which is modeled as a function from the logical space of possible worlds to a chronology, which in turn is modeled as a (partial) function from a domain of times to entities such as individuals, numbers, sets, etc. Programmatically stated, Tichý takes Frege's semantics for oblique ('*ungerade*') contexts and universalizes it so as to apply to all contexts. One important addition is the implementation of empirical indices such as worlds and times to model both modal and temporal variability. Another important addition is that it is insufficient to have 'the F' denote a *Sinn* or intension without the option to descend from the *Sinn* or intension to what it presents or has as a value (at the index pair of evaluation). We want to be spared the embarrassment that a *Sinn* or intension is a celestial body illuminated by the Sun. Simply elevating the semantic value from a planet to a mode of presentation of a planet, as Frege does, is only half the job. Tichý accounts for the extensionalization of an intension by way of the intension being applied to the empirical indices of evaluation: given a world and a time, we are given the entity (if any) who or which occupies the office of the F at that dual index. Similarly, predicates do not denote, or have as their semantic value, either a set or a multitude of individuals, but a property; sentences do not denote a truth-value, but a truth-condition, which is what possible-world semantics knows as a 'proposition'.[20]

According to Tichý, 'Nancy Reagan's spouse' never has as its semantic value, or denotation, Ronald Reagan or any other individual. Rather its semantic value is the *condition* of being Nancy Reagan's spouse. Similarly, the semantic value of 'The Morning Star' is not Venus or any other individual, but the condition of being the brightest non-lunar celestial body in the morning sky. Reference shift has been abolished. Consequently and crucially, the reason 'Hesperus', 'Phosphorus' (or rather their respective denotations) do not substitute is not because they would be identicals that opacity prevented from being substituted. There is no such thing as

[19]Frege, we might add, would also have trouble with this variant of quantifying-in: $\exists x(Fx \wedge Bel\ a\ (\ldots x \ldots))$. The two occurrences of x seem not to share the same value, the two values presumably being an extension and an intension, according to reference-shift theories. If they differ, an attempt to quantify-in will bring out this incongruity: there is no value of x such that that value is both an F and believed by a to be such-and-such.

[20]Tichý models a *property* as a function from worlds to a function from times to sets of individuals, identifying a set with its characteristic function, and he models a *proposition* as a function from worlds to a partial function from times to truth-values (i.e. truth-value gaps are an option).

opacity in Tichý's semantics. Instead the reason is because they are not identicals. So the reason why Leibniz's Law 'fails' is because it is being misapplied to the wrong substituends, namely two non-synonymous, indeed not even logically equivalent definite descriptions (or offices/intensions, in the material mode).

So how are those two conditions, of being Nancy Reagan's spouse and being the brightest non-lunar celestial body in the morning sky, to be brought into contact with Reagan and Venus? By means of an additional premise, which states an empirical fact:

(i.iv) At $\langle w, t\rangle$ Nancy Reagan's spouse is male
(ii.iv) At $\langle w, t\rangle$ Nancy Reagan's spouse is Ronald Reagan

---

(iii.iv) At $\langle w, t\rangle$ Ronald Reagan is male

(i.v) At $\langle w, t\rangle$ the Morning Star is a planet
(ii.v) At $\langle w, t\rangle$ the Morning Star is Venus

(iii.v) At $\langle w, t\rangle$ Venus is a planet

Premises (i.iv/v) share the general linguistic form "The *F* is a *G*". Tichý offers various ways of logically analyzing this form, depending on the degree of semantic structure we wish to make explicit. For instance, (i.iv) obtains an individual (the person married to Nancy Reagan, or else nobody) via another individual (Nancy Reagan), while (ii.iv) obtains an individual (or else nobody) in one go. These differences are less relevant here, for all I want to make a case for at this point is that empirical definite descriptions denote a condition rather than the satisfier, if any, at the index or indices of evaluation. Kaplan, like so many before and after him, assumes that the right sort of substituends must be a pair of definite descriptions as soon as they happen to share the same descriptum, the background assumption being that the semantic value of a definite description is its descriptum. This overly extensionalist theory of definite descriptions is bound to run into trouble amidst non-extensional contexts, with referential opacity looming on the horizon.

In Sect. 12.3 I will present all the relevant semantic, including type-theoretic, details. At this juncture I will present the basic ideas in a semi-formal way to get the philosophy across. In case we construe the definite descriptions 'Nancy Reagan's spouse' and 'The Morning Star' as having identical semantic structure, then (i.iv/v) share this logical form (to be revised in Sect. 12.3), which underlies "The Morning Star is a planet"[21]:

$$\lambda w \lambda t[Planet_{wt} MS_{wt}]$$

Let me decompose this complex, structured whole into its constituent parts, followed by a type assignment. I am treating, for now, the modal and temporal

[21] See Tichý [27], Duží [8].

indices as if they formed one index, thus glossing over subtleties that are irrelevant here.

- $Planet$: property
- $Planet_{wt}$: set
- MS : office
- MS_{wt} : individual
- $[Planet_{wt} MS_{wt}]$: (a presentation of a) truth-value
- $\lambda w \lambda t[Planet_{wt} MS_{wt}]$: (a presentation of a) truth-condition (proposition)

$[Planet_{wt} MS_{wt}]$ is the application of a set to an individual with a view to obtaining a truth-value. If MS is undefined at $\langle w, t\rangle$, the application fails and no truth-value emerges. If MS is defined at $\langle w, t\rangle$, the application of $Planet_{wt}$ to MS_{wt} yields a truth-value, according as the individual that is the extension of MS at $\langle w, t\rangle$ is a member of the extension of $Planet$ at $\langle w, t\rangle$. The emerging truth-value is abstracted over by λt, and the resulting chronology (here, a function from times to truth-values) is abstracted over by λw.[22] An empirical truth-condition is, allowing a slight simplification, the set of world-time pairs at which a given function from worlds and times to truth-values returns True. In set-theoretic terms, the truth-condition is this:

$$MS\langle w, t\rangle \in Planet\langle w, t\rangle$$

It is an open empirical question whether the actual world and the present time are members of the satisfaction class of this proposition.[23] Tichý's semantics is opposed to privileging the actual world and the present time, treating instead all worlds and all times as being equal.[24]

(ii.iv/v) also share the same form, in case we align the former with the latter. Thus the semantic structure of "The Morning Star is Venus" is:

$$\lambda w \lambda t[= MS_{wt} Venus]$$

The newcomers are these two:

- Venus : individual
- = : identity between individuals

[22]Hence it is a slight simplification when I use the notation '$\langle w, t\rangle$' as though intensions were defined on a *pair* of arguments rather than being defined on, first, worlds and, next, on times. The simplification is innocuous enough in this essay, but it is crucial to be technically able to treat the modal and the temporal separately. Tichý [30] explains why, and so does Duží et al. [8, pp. 205–27], offering more by way of technical and philosophical exposition.

[23]The present time is here both the time *at* which and *about* which it is being asserted that the Morning Star is a planet. Hence the index of assertion and the index of evaluation coincide here.

[24]See Tichý [27] and Duží et al. [8, pp. 178–90].

Identity relations, according to Tichý, are polymorphous functions which, when type-theoretically well-defined, take a pair of objects of the same type to a truth-value, according as the two functional arguments are one and the same object.[25] The challenge, then, is to capture the contingency of the truth (or falsity or truth-value gap, for that matter) that the Morning Star is Venus without infringing the absolute self-identity of Venus. First and foremost, the Morning Star and Venus are two distinct entities, even of two distinct types (an office and an individual, respectively). So what "The Morning Star is Venus" means is that, at $\langle w, t \rangle$, the office of Morning Star is occupied by Venus. The condition is that whatever individual is the occupant of *MS* at $\langle w, t \rangle$ be identical to Venus; or equivalently, due to the symmetry of identity, that Venus be identical to whatever individual is the occupant of *MS* at $\langle w, t \rangle$.

Notice that Tichý construes "The Morning Star is Venus", "The Morning Star is a planet", and all the other sentences we brought up above, as *modal* contexts. They all have a modal profile, and it is the same for all of them, namely contingency. Thus, a more careful analysis would rephrase, e.g., "The Morning Star is Venus" as "Contingently, the Morning Star is Venus". This is important, for this implies that every empirical sentence is an intensional context. Tichý does not introduce a contingency operator; rather he uses abstraction over at least worlds and often also times to capture contingency.[26] To amplify the point made above, the more careful analysis is *not* "Actually and presently, the Morning Star is Venus".

In the full, final analysis presented in Sect. 12.3, ${}^0MS_{wt}$ is a three-step *procedure* that is *typed* to produce or present an individual. Whether the procedure, if executed, does indeed produce or present an individual—namely the individual that occupies the office of Morning Star at the $\langle w, t \rangle$ of evaluation, and if so, which individual is thus produced—is beyond the procedure. Similarly, ${}^0Planet_{wt}$ is a three-step procedure that is typed to produce a set of individuals at $\langle w, t \rangle$. There is always going to be a set of planets, although it may be the empty set of individuals. In Sect. 12.3 the notation will become slightly, but importantly, more elaborate, because we want to identify the various parts of a whole, each part being a sub-procedure within a complex procedure. This is important because we need to gain access to some of those parts inside a *displayed* (as opposed to *executed*) procedure in order to manipulate them. In Tichý's theory, whenever an agent entertains a hyperpropositional attitude, the agent is related intentionally to a displayed

[25]See Duží et al. [8, pp. 296–300] for a survey of different sorts of identity statements, and (ibid., pp. 301–10) for a detailed analysis of "Hesperus is Phosphorus", predicated on the premise that 'Hesperus', 'Phosphorus' denote two distinct offices. The sentence does not express the necessary self-identity of an individual bearing two names, but the contingent convergence of two distinct, differently named conditions in the same anonymous satisfier.

[26]An obvious exception to the standard pattern of $\lambda w \lambda t[\ldots w \ldots t \ldots]$ is the modeling of nomological ('soft') necessity, which is captured thus: $\lambda w \forall t[\ldots w \ldots t \ldots]$. Relative to a set of worlds, such-and-such is always or never true, i.e. laws of nature are necessary and atemporal only relative to an equivalence class of nomologically indistinguishable worlds. The major source of inspiration is the Armstrong-Dretske-Tooley conception of nomological necessity. See Duží et al. [8, pp. 411–14].

procedure, where the attitude complement is the very procedure rather than what it produces when executed.[27] That is to say, in Fregean parlance, that the agent is related to a mode of presentation of a truth-condition. The twist is that the agent may well be innocent of which truth-condition is so presented.

12.2 Two Arguments Against Unqualified Quantifying-In

The general line of objection both to modalities and attitudes consists in arriving at a paradoxical property—like being necessarily, and also not necessarily, two-legged; or being believed, and also not being believed, by *a* to be a spy. In the light of paradox, quantifying-in becomes a spurious undertaking, for how could there be any value of *x* with this or that paradoxical property?

If existential quantification is not an option then, ipso facto, one of the hallmarks of extensional logic is not an option. Hence something somewhere is not quite right. Is the culprit a poor analysis of one or more of the premises, or is it any of the inference rules? Let me spill the beans straightaway. If you ask me, the arguments are flawed because of a deficient analysis of singular terms such as definite descriptions, which leads to a deficient analysis of identity sentences flanked by at least one definite description, which leads to wrongful applications of rules of extensional logic. The definite descriptions 'The man on the beach', 'The man in the brown hat' may contingently share the same unique descriptum (an individual). If they do, they still do not denote, or are in any semantically significant way 'about', their shared descriptum. Rather, to restate the claim, they denote two distinct empirical conditions. The logical form of "The man on the beach is the man in the brown hat" is that two distinct conditions, named 'The man on the beach' and 'The man in the brown hat', contingently share the same anonymous satisfier. Hence, despite the contingent convergence of the two conditions, it is one thing to know or believe that the man on the beach is a spy and quite another that the man in the brown hat is a spy. Hence *a* does not entertain the contradictory belief that an individual both is, and is not, a spy. Hence it does not follow that there is an individual *x* such that *a* believes, and does not believe, that *x* is a spy.

The key move, thus, is to bar Leibniz's Law from applying, rather than dismissing the Law as invalid. The Law is valid, for sure, also in the most exacting attitude contexts, but it is not always applicable. Its substituends must be chosen carefully. A mere coincidence of two conditions in one satisfier is not good enough for substitution inside an attitude context. What is substitutable are hyperintensions for hyperintensions, intensions (functions) for intensions, and extensions (functional values) for extensions.

[27] Tichý's original term was 'constructional attitude': see his [31, pp. 221–24]. 'Construction' is the term Tichý coined for his structured hyperintensions: see Sect. 12.3 of this essay. Tichý never got around to developing his notion of constructional/hyperintensional attitude to any great degree.

It is also important to keep the semantic and the logical issues apart while keeping their interplay in mind. Issues like quantification and substitution of identicals are logical ones; issues like the sense and denotation of definite descriptions and the sense of identity sentences flanked by definite descriptions are semantic ones. The interplay is first and foremost that a proper semantic analysis of 'The *F*'/'The *G*' and "The *F* = the *G*" must precede the application of Leibniz's Law and EG. Faced with an insufficient semantic analysis, the Law and/or EG are liable to being misapplied.

12.2.1 The Argument from Non-factivity

The idea behind this particular argument is easily stated. An instance of EG will take us from a truth to a falsehood if EG conjures an object—an arbitrary value of an ∃-bound variable—into existence. Here is an example:

$$(5)\ \frac{\text{Tilman is seeking the fountain of youth}}{\exists x\ (\text{Tilman is seeking } x)}\ \text{EG}$$

where *x* ranges over *individuals*, on the assumption that the fountain of youth may be construed metaphysically as an individual. (5) violates the constraint that we may well seek what fails to exist. What fails to exist may still be sought; seekers may fail to be finders. For a non-empirical example:

$$(6)\ \frac{\text{Tilman is calculating the quotient of dividing 5 by 0}}{\exists y\ (\text{Tilman is calculating } y)}\ \text{EG}$$

where *y* ranges over *natural numbers*. EG is not a magic wand we can wave to create a number where there was previously none. What is more, I am not comfortable with having Tilman be computationally related to a number. Calculation, in a nutshell, is all about applying operations to numbers in their capacity as operands in order to obtain a new number. And the computational effort may be futile, either because the agent lacks the sufficient skill to complete the calculation, or because the operation and one or more of the operands are a mismatch. Tilman's arithmetic predicament is that 0 is not a suitable divisor. The morale I draw is that this arithmetic predicament must not affect the semantics of Tilman's calculating the quotient of $\langle 5, 0\rangle$. The sentence "Tilman is calculating the quotient of dividing 5 by 0" is perfectly meaningful and an apt vehicle for making an assertion about what Tilman is up to. So what we do is take a step back. Tilman is to be related computationally to a procedure and not the sort of object the procedure is typed to produce. It does affect the logic of Tilman's activity that 0 is a dysfunctional divisor. For his attitude is non-factive, hence quantifying over numbers is not an option, hence quantifying-in seems to be impossible, hence the attitude reported in the premise seems somewhat iffy. But Tilman's attitude is just fine, and quantifying-in is valid. We just need to quantify over procedures instead. The analysis of (6) actually proves helpful in unearthing the nature of Tilman's complement: where we might naïvely have expected a number, we realize we need instead a procedure typed to produce numbers. The format of the proper analysis, I submit, is (6'):

$$\text{(6')}\ \frac{\text{Tilman is calculating the quotient of dividing 5 by 0}}{\exists c\ (\text{Tilman is calculating } c)}\ \text{EG}$$

What is going on in the premise is that Tilman is intentionally related to some procedure that he expects to deliver the quotient of dividing 5 by 0. The premise does not presuppose this or that particular procedure, e.g. the procedure of applying the division function to $\langle 5, 0 \rangle$. The conclusion does nothing other than spell out that there is some procedure or other that Tilman is deploying to this end.

(5) is a different kettle of fish by opening up the prospect of also quantifying over intensions.[28] Of course, there is also going to be a procedure (multiple, in fact) producing a given intension, so we can quantify over procedures as well, but quantifying over the lower-level entities of intensions makes for a stronger conclusion. What is going on in the premise is that Tilman is in the process of tracking down the occupant of an office. It is the office of the fountain of youth that is guiding Tilman's search and is what he is intentionally related to. The format of the proper analysis, I submit, is (5'):

$$\text{(5')}\ \frac{\text{Tilman is seeking the fountain of youth}}{\exists f (\text{Tilman is seeking } f)}\ \text{EG}$$

where f ranges over offices. (5') is a case of quantifying into an intensional context and quantifying over intensions of a particular type (offices, as it happens). It is an additional option to also quantify over hyperintensions producing intensions. (6') is a case of quantifying into a hyperintensional context *and* quantifying over hyperintensions of a particular type (those producing natural numbers, as it happens).

In general, the argument from non-factive attitudes against quantifying-in has bite only if we insist on quantifying into a non-extensional context and quantifying over extensions such as individuals and numbers. The argument is barking up the wrong tree as soon as we include intensions and hyperintension (procedures) into our ontology for our variables to quantify over. The retreat, as it were, or ascent to non-extensional entities is not an attempt to dodge the problem. Rather it serves to get clearer about the attitude complement in the premise.

[28] Zimmermann [33, p. 728] brings up this 'unwelcome inference':

$$\frac{\text{I owe you nothing [i.e. there is nothing I owe you]}}{\text{I owe you something [i.e. there is something I owe you]}}$$

Unwelcome it is, in case we are quantifying over extensional entities. But suppose we are quantifying over *properties*:

$$\frac{\text{There is no property of which I owe you an instance}}{\text{There is at least one property of which I do not owe you an instance}}$$

Then we obtain a valid argument favouring quantifying-in.

12.2.2 The Argument from Doublethink

Quine [20] has us consider these three assumptions:

(1) Ralph believes that the *F* is an *H*
(2) Not: Ralph believes that the *G* is an *H*
(3) The *F* = the *G*

---

(4) ?

Quine then asks: does Ralph, or does he not, believe that the individual who is both the *F* and the *G* is an *H*? In Quine's example, Ralph believes that the man he has seen in the brown hat is a spy; Ralph does not believe (i.e. abstains from believing) that the man he has seen on the beach is a spy; the man in the brown hat and the man on the beach are one and the same man.

The tricky bit is not that it is inferable that there is an individual x such that Ralph has both seen x wear a brown hat and has seen x on the beach. The tricky bit is rather that it seems we could infer that there is an individual x such that Ralph both believes, and does not believe, that x is a spy. What quantification does is quantify away the two different aspects under which Ralph has made acquaintance with an individual, retaining only the individual himself, in his capacity as a value of x. Quine's question is unreasonable, in my view, for Ralph does not know that (3) is true. If he did, he would rationally have to update his belief base. Only then would it make sense for Quine to ask his question. The reason, of course, why Quine raises his question is because he wants to challenge the logic underlying (1) and (2). From the above premise set it seems we can extract the contradiction that there is an x such that Ralph believes that x is a spy and Ralph does not believe that x is a spy.

A *reductio* enjoins us to reconsider the premises and the rules of inference. The solution, as suggested above, is to make the semantics of 'the *F*'/'the *G*' a pivotal point. What we can infer from (1) is that there is an office y such that Ralph believes that the occupant of y at $\langle w, t\rangle$, i.e. y_{wt}, is a spy. What we can infer from (2) is that there is an office z such that Ralph believes that z_{wt} is a spy. What we can infer from (3) is that there are two offices z', z'' such that $z'_{wt} = z''_{wt}$, i.e. that z', z'' are co-occupied at $\langle w, t\rangle$. What we cannot infer from (1) through (3) is that there is an office z''' such that Ralph believes that z'''_{wt} is a spy and Ralph does not believe that z'''_{wt} is a spy. The reason for this is that z'_{wt}, z''_{wt} are not interchangeable inside (1), (2). These contexts are not extensional, for what matters is not the convergence between z'_{wt}, z''_{wt} in some anonymous individual, as reported by (3). Instead what matters is whether 'z'', 'z''' co-denote the same office. And they do not. So 'z'_{wt}' cannot be substituted for 'z''_{wt}' in the context "…believes…z''…", or the other way around. The upshot is that (3) becomes irrelevant. With (3) irrelevant, only (1) and (2) matter, and they are too weak to sustain the contradictory conclusion we have been trying to avoid all along.

12.3 Validating Quantifying-In

In this section I present the logical details required to pull off quantifying into a non-factive, hyperpropositional attitude context by means of Tichý's framework called Transparent Intensional Logic. Formally, TIL is a partial, typed, hyperintensional λ-calculus. Its λ-terms are interpreted by way of a procedural (as opposed to denotational or extensionalist) semantics. TIL belongs squarely to the paradigm of hyperintensionally individuated, structured meaning. TIL is arguably the most elaborate procedural semantics for logical analysis of natural language on the market.[29] TIL arguably also offers what are presently the most worked-out criteria of hyperintensional individuation.[30] What is of primary importance here is that TIL has an elaborate theory of hyperintensional attitudes.

Still new theories of hyperintensions are being spawned at present. They divide into at least three different kinds: (i) hyperintensions as primitive black-boxes (e.g. Thomason [26], Bealer [1] and later); (ii) Cresswell-style hyperintensions as set-theoretic sequences; and hyperintensions as procedures (e.g. Hanks [12, 13], Soames [24]). My [14] explains why I have little time for (i) and (ii), and why I much prefer (iii). The basic idea is that hyperintensional individuation is procedural individuation, and that procedures are very apt at accommodating very fine semantic and logical differences. For instance, a procedural semantics worth its name will assign two distinct meanings to '7 + 5' and '5 + 7'. Thus one thing is to calculate the sum of $\langle 7, 5\rangle$, and quite another to calculate the sum of $\langle 5, 7\rangle$, because the respective arguments are organized in two distinct sequences.

The example I will be analyzing here is the following sentence:

"Le Verrier believes that Vulcan causes Mercury's perturbations" (*analysandum*)

Le Verrier had discovered Neptune in 1846, his observation of the planet being preceded by mathematical calculations and speculations. He claimed, in 1860, to have established the existence of a planet orbiting between Mercury and the Sun, also on the basis of observation (though somebody else's) preceded by calculation and speculation. The grammatical proper name 'Vulcan' (or 'Vulcanus'), originally used for the Roman god of fire and volcanoes, seemed an apt name for a planet so close to the Sun. Problem is, no planet has so far ever been found between Mercury and the Sun. Must the semanticist draw the conclusion that the astronomer's 'Vulcan' lacks a semantic value? The conclusion seems tempting; but it is a temptation that should not be yielded to. For Le Verrier has a belief *about* something,

[29] See Duží et al. [8, Ch. 1].

[30] Duží and Jespersen [7, Sec. 2] motivates and defines TIL's latest criterion, called (A"), which is a variant of Church's Alternative (A1), by containing a slightly more detailed definition of α-conversion and β-conversion-by-value, whereas Church has α-conversion and β-conversion-by-name. The main reason we are leaving out η-conversion is that it fails to guarantee equivalent conversion in a partial logic such as TIL. See also Jespersen [15, Section 3, Def. 4] (2015, Section 3, Def. 4) for the exact definition of A".

and his intentional act is one I want to take seriously in the sense of according an abstract object to it. What we can rule out is that his intention was directed at an *individual*, for there is none at the receiving end, unlike what holds for his beliefs about Neptune. So this attempt to quantify-in will be a wheel spinning in the void[31]:

There is an individual x such that Le Verrier believes that x causes Mercury's perturbations.

But this one latches on to an entity:

There is an office f such that Le Verrier believes that the occupant of f causes Mercury's perturbations.

So does this one:

There is a hyperintension h such that Le Verrier believes that the occupant of the office constructed by h causes Mercury's perturbations.

I stipulate Le Verrier's doxastic attitude to be a relation-in-intension between an individual and a hyperproposition. Alternatively, his complement might have been an intension, *in casu* a proposition. But my reason for going with the hyperproposition is to demonstrate how to *quantify into a hyperpropositional attitude context* and make the strongest quantification we can make when non-factive attitudes are involved, which is in this case *quantification over offices.*

12.3.1 Basic Definitions

To proceed from the philosophical remarks made above to a logical theory of quantifying-in, we need to introduce the following basic definitions, which also introduce the relevant notation. The notation will be implemented in Sect. 12.3.2. Section 12.3.3 presents and proves the validity of the rule of quantifying-in that our analysandum requires.

Definition 12.1 (types of order 1). Let B be a *base*, where a base is a collection of pair-wise disjoint, non-empty sets. Then:

[31] One might still wonder why that conclusion does not follow. The reason is that this existential quantification takes place *outside* Le Verrier's doxastic perspective. The actual celestial scheme of things does not include an individual that causes Mercury's perturbations. Hence the conclusion would be false. This becomes especially clear if we phrase the conclusion as a belief *de re*: There is an individual x of which Le Verrier believes that it causes Mercury's perturbations; or: There is an individual x that is believed by Le Verrier to cause Mercury's perturbations (cf. n.35). This sort of attribution would be forthcoming if, for instance, Le Verrier had first identified Venus and then gone on to misidentify Venus as the cause of Mercury's perturbations. The opposite case, with the quantifier inside the scope of the belief operator, would state an uncontroversial truth: Le Verrier believes that there is an individual x such that x causes Mercury's perturbations. (Thanks to Alessandro Torza for raising the initial point.)

(i) Every member of B is an elementary *type of order 1 over B*.
(ii) Let $\alpha, \beta_1, \ldots, \beta_m (m > 0)$ be types of order 1 over B. Then the collection $(\alpha\beta_1 \ldots \beta_m)$ of all m-ary partial mappings from $\beta_1 \times \ldots \times \beta_m$ into α is a functional type of order 1 over B.
(iii) Nothing else is a *type of order 1 over B*. □

For the purposes of natural-language analysis, we are currently assuming the following base of *ground types*, each of which is part of the ontological commitments of TIL:

o: the set of truth-values $\{\mathbf{T}, \mathbf{F}\}$;
ι: the set of individuals (constant universe of discourse);
τ: the set of real numbers (doubling as temporal continuum);
ω: the set of logically possible worlds (logical space).

Construction is Tichý's formal notion of structured hyperintension, interpreted philosophically as a procedure.[32]

Definition 12.2 (construction).

(i) (*Variable*) Let valuation v assign object o to variable x. Then x is a *v-construction of o*.
(ii) (*Trivialization*) Let X be any object whatsoever (i.e. an extension, an intension, or a *construction*). Then ${}^0X$ is the *Trivialization of X*, which *constructs X* without any change of X.
(iii) (*Composition*) Let $X, Y_1, \ldots Y_m$ be objects of any type. Let X *v-construct* a function f of type $(\alpha\beta_1 \ldots \beta_m)$, and let $Y_1, \ldots, Y_m$ *v-construct* entities $B_1, \ldots, B_m$ of types $\beta_1, \ldots, \beta_m$, respectively. Then the *Composition* $[XY_1 \ldots Y_m]$ *v-constructs* the value (an entity, if any, of type α) of f on the tuple argument $\langle B_1, \ldots, B_m \rangle$. Otherwise the *Composition* $[XY_1 \ldots Y_m]$ does not *v-construct* anything and so is *v-improper.*
(iv) (*Closure*) Let $x_1, \ldots, x_m, Y$ be objects of any type. Let $x_1, \ldots, x_m$ be pairwise distinct variables *v-constructing* entities of types $\beta_1, \ldots, \beta_m$, and let Y be a *construction v-constructing* an α-entity. Then $[\lambda x_1 \ldots x_m Y]$ is the *construction λ-Closure* (or *Closure*). It *v-constructs* the following function f of type $(\alpha\beta_1 \ldots \beta_m)$. Let $v'(B_1/x_1, \ldots, B_m/x_m)$ be a valuation identical with v at least up to assigning objects $B_1/\beta_1, \ldots, B_m/\beta_m$ to variables $x_1, \ldots, x_m$. If Y is $v'(B_1/x_1, \ldots, B_m/x_m)$-improper (see iii), then f is undefined on $\langle B_1, \ldots, B_m \rangle$. Otherwise the value of f on $\langle B_1, \ldots, B_m \rangle$ is the α-entity $v'(B_1/x_1, \ldots, B_m/x_m)$-*constructed* by Y.
(v) Nothing else is a *construction*. □

Here are some informal explications of each kind of construction.[33] Bear in mind that the overarching idea behind the notion of construction is that, given some input

[32] I am leaving out two constructions—*Single* and *Double Execution*—which I do not need when quantifying into the particular hyperintensional contexts considered here.

[33] These explications draw on material from Jespersen [15, Section 3].

objects, we can apply operations or procedures or *constructions* to obtain some output objects (or none, in some instances of Composition). A *variable* constructs an object by having this object as its value dependently on a valuation function v arranging variables and objects in a sequence. *Trivialization* is TIL's objectual or material counterpart of a non-descriptive constant term, which simply harpoons a particular object. *Composition* is the procedure of functional application, rather than the functional value (if any) resulting from application. *Closure* is the procedure of functional abstraction, rather than the resulting function.

Variables and Trivializations are the one-step or primitive or *atomic* constructions of TIL. In particular, what does not exist cannot be Trivialized. (Similarly, what does not exist cannot be named; but it can be described, as per 'the largest prime' or 'is a winged unicorn'.) Composition and Closure are the multiple-step or *composite* procedures. An *atomic* construction is a structured whole with but one proper part, namely the construction itself. Importantly, the proper part of ^0X is ^0X and not X, which is located beyond ^0X: the product of a procedure is no part of the procedure. A *composite* construction is a structured whole with more proper parts than just itself.

The definition of the typed universe of TIL amounts to a definition of the ramified hierarchy of types which divides into three parts; firstly, simple types of order 1, which were already defined by Definition 12.1; secondly, types belonging to order $*_n$; thirdly, types of order $*_{n+1}$.

Definition 12.3 (ramified hierarchy of types).

T_1 *(types of order 1)*. See Definition 12.1

C_n *(constructions of order n)*.

(i) Let x be a variable ranging over a type of order n. Then x is a *construction of order n over B*.
(ii) Let X be a member of a type of order n. Then ^0X is a *construction of order n over B*.
(iii) Let $X, X_1, \ldots, X_m$ $(m > 0)$ be constructions of order n over B. Then $[XX_1 \ldots X_m]$ is a *construction of order n over B*.
(iv) Let $x_1, \ldots x_m, X$ $(m > 0)$ be constructions of order n over B. Then $[\lambda x_1 \ldots x_m X]$ is a construction of order n over B.
(v) Nothing is a *construction of order n over B* unless it so follows from C_n (i)–(iv).

T_{n+1} (types of order $*_{n+1}$). Let $*_n$ be the collection of all constructions of order $*_n$ over B. Then

(i) $*_n$ and every type of order n are *types of order n+1*.
(ii) If $m > 0$ and $\alpha, \beta_1, \ldots, \beta_m$ are types of order $n+1$ over B, then $(\alpha\beta_1 \ldots \beta_m)$ (see T_1 ii) is a *type of order n + 1 over B*.
(iii) Nothing else is a *type of order n + 1 over B*. □

As a notational convention, 'X/α' means that object X is of type α.

Definition 12.4 (existential quantifier). The existential quantifier $\exists\alpha$ is a type-theoretically polymorphous, total function of type $(o(o\alpha))$, for an arbitrary type α, defined as follows: $\exists\alpha$ is the function that associates a class S of α-elements with T if S is a non-empty class, otherwise with F. □

An occurrence of a bound variable is either ^{0}-bound ('Trivialization-bound') or λ-bound or both. Bound occurrences contrast with free occurrences. Hence:

Definition 12.5 (free and bound occurrences of variables). Let C be a construction with at least one occurrence of a variable ξ.

(i) Let C be ξ. Then the *occurrence of* ξ in C is free.
(ii) Let C be ^0X. Then every *occurrence of* ξ in C is ^{0}*-bound.*
(iii) Let C be $[\lambda x_1 \ldots x_n Y]$. Any *occurrence of* ξ in Y that is one of x_i, $1 \leq i \leq n$, is λ*-bound* in C unless it is ^{0}*-bound in Y.* Any *occurrence of* ξ *in* Y that is neither ^{0}*-bound* nor λ*-bound* in Y is *free* in C.
(iv) No other *occurrence* of ξ is *free,* λ*-bound,* ^{0}*-bound* in C. □

12.3.2 *Executed Versus Displayed Occurrences of Constructions, and Two Additional Functions*

If our analysandum above were an intensional attitude, as per possible-world semantics, then the analysis would be this, where P is the relational property, of type $(o\iota)_{\tau\omega}$, of causing Mercury's perturbations:

$$\lambda w\lambda t[{}^0B^*_{wt}\,{}^0LeV\lambda w\lambda t[{}^0P^0_{wt}V_{wt}]]$$

B is typed as a relation-in-intension between an individual (a doxastic agent) and a proposition (truth-condition): $B/(o\iota o_{\tau\omega})_{\tau\omega}$.

The hyperproposition $\lambda w\lambda t[{}^0P_{wt}{}^0V_{wt}]$, of type $*_1$, produces a proposition, of type $o_{\tau\omega}$, and Le Verrier believes that this proposition is true. TIL says that $\lambda w\lambda t[{}^0P_{wt}{}^0V_{wt}]$ occurs *executed* in the context $\lambda w\lambda t[{}^0B^*_{wt}\,{}^0LeV\lambda w\lambda t[{}^0P_{wt}{}^0V_{wt}]]$, because it descends to its product (a proposition). Le Verrier's attitude becomes a hyperintensional one as soon as $\lambda w\lambda t[{}^0P_{wt}{}^0V_{wt}]$ occurs *displayed*.[34] The hyperintensional analysis is this:

$$\lambda w\lambda t[{}^0B^*_{wt}\,{}^0LeV{}^0[\lambda w\lambda t[{}^0P_{wt}{}^0V_{wt}]]]$$

B^* is typed as a relation-in-intension between an individual (a doxastic agent) and a hyperproposition: $B^*/(o\iota *_1)_{\tau\omega}$. With Le Verrier as the first argument of B^*, the

[34] Duží and Jespersen [7, Sections 2.4 and 3.2] offers a detailed exposition of the twin notion of executed and displayed occurrence of constructions.

second argument is the construction $[\lambda w \lambda t[^0P^0_{wt}V_{wt}]]$. What Le Verrier believes* is that the hyperproposition presents a proposition that is true.

The advertised qualifications of quantifying-in are the following two. First, B/B^* is a non-factive attitude, so we need to be careful when selecting a quantificational range for the ∃-bound variable x occurring in the hyperintensional context

$$^0[\lambda w \lambda t[^0P_{wt}x]]$$

The range cannot be ι; I suggest $\iota_{\tau\omega}$, i.e. offices occupiable by individuals. If we go along with this typing, we need to extensionalize the value of x to obtain the right type of object of which to predicate the property of being a P. Hence

$$^0[\lambda w \lambda t[^0P_{wt}x_{wt}]]$$

Second, x will be *doubly bound*. When saying that x is ∃-bound I am being imprecise, for we are inside a λ-calculus, so in principle all variable-binding is λ-binding. Hence $\exists\lambda x$ is what we want: the function ∃ is applied to the set produced by λx, yielding a truth-value (just as in Montague). However, TIL has an additional form of variable-binding, introduced in Definition 12.5 (ii), namely so-called Trivialization-binding: ^0x. Thus x is both λ-bound and Trivialization-bound—and the latter trumps the former, for the Trivialization of a context makes each of its constituent parts occur Trivialized as well, as per Definition 12.4 (ii):

$$^0\exists\lambda x \ldots {}^0[\lambda w \lambda t[^0P_{wt}x_{wt}]]$$

This means that x occurs displayed in $^0[\lambda w \lambda t[^0F_{wt}x_{wt}]]$, for each component part of this context occurs displayed, without x descending to any of the values in its quantificational range. The upshot is that the λ of $\exists\lambda x$ cannot bind x. For this λ to bind x, x would have to occur executed in order to descend to a value. This explains why the following will not do:

$$\lambda w \lambda t[^0\exists\lambda x[^0B^*_{wt}{}^0LeV^0[\lambda w \lambda t[^0P_{wt}x_{wt}]]]]$$

This simplistic analysis of quantifying-in would be not unlike the futile attempt to quantify into a quoted context, an undertaking Quine would frequently, fervently and also correctly warn against. Without suitable qualification, quantifying-in will be rendered illogical. We need a much more subtle way in than forced entry. The logical challenge is to make a Trivialization-bound occurrence of x amenable to λ-binding. Some logical work needs to be performed in the '?'-marked area of $^0\exists\lambda x \ldots ? \ldots {}^0[\lambda w \lambda t[^0P_{wt}x_{wt}]]$.

TIL makes x amenable to λ-binding by means of the polymorphous functions *Sub* and *Tr*. *Sub* substitutes (all the occurrences of) one construction for (all the occurrences of) another construction inside a third construction to obtain a fourth construction. Substitution is construed as a primitive operation. *Tr* takes an entity of the arbitrary type α and returns its Trivialization.[35]

Definition 12.6 (*Sub^n*). Let $C_1/*_{n+1} \rightarrow *_n$, $C_2/*_{n+1} \rightarrow *_n$, $C3/*_{n+1} \rightarrow *_n$ v-construct constructions D_1, D_2, D_3, respectively. Then the Composition $[{}^0Sub^n C_1 C_2 C_3]$ v-constructs the construction D that results from D_3 by collisionless substitution of D_1 for all occurrences of D_2 in D_3. □

Definition 12.7 (*Tr^α*). The function $Tr^\alpha/(*_n\alpha)$ returns as its value the Trivialization of its α-argument. □

For instance, let variable y v-construct entities of type ι, such as a. Then $[{}^0Tr^\iota y]$ $v(a/y)$-constructs ${}^0a$. Therefore, the Composition $[{}^0Sub^1[{}^0Tr^\iota y]{}^0x^0[{}^0F_{wt}x]]$ $v(a/y)$-constructs the Composition $[{}^0F_{wt}{}^0a]$. Note that there is a substantial difference between the *construction* Trivialization and the *function* Tr^α. Whereas ${}^0y$ constructs just the variable y regardless of valuation, y being ^{o}-bound in ${}^0y$, $[{}^0Tr^\iota y]$ v-constructs the Trivialization of the object v-constructed by y. Hence y occurs free in $[{}^0Tr^\iota y]$.

12.3.3 Rules of Quantifying-in

We now have everything we need to introduce the *rule of existential quantification into hyperpropositional attitude contexts*. For the first rule, let the type-theoretic assignments be as follows, where '$C \rightarrow_v \alpha$' means that construction C constructs an entity of type α dependently on a valuation v of the variables involved: as above, $B^*/(o\iota *_n)_{\tau\omega}$ is a hyperpropositional attitude relation; $\exists/(o(o*_n))$; a an individual of type ι; $C(X)/*_n \rightarrow_v o_{\tau\omega}$ a propositional construction with a constituent $X/*_n$ such that X v-constructs entities of type α; $d/^*_n \rightarrow_v \alpha$; $c/*_{(n+1)} \rightarrow_v *_n$. This last type assignment makes it clear that quantifying into hyperintensional contexts requires a fair amount of expressive power, since we need a higher-order construction to construct a lower-order construction.

$$\frac{[B^*_{wt}\ {}^0a{}^0C(X/d)]}{[{}^0\exists\lambda c[B^*_{wt}\ {}^0a[{}^0Sub\ c\ {}^0d\ {}^0C(d)]]]}\ \textit{Rule A}$$

The basic idea behind the proof of the rule is as follows. The Composition $[{}^0Sub\ c\ {}^0d\ {}^0C(d)]$ $v(X/c)$-constructs the construction $C(X)$. Hence at any $\langle w, t\rangle$ at which $[B^*_{wt}\ {}^0a\ {}^0C(X/d)]$ v-construct **T**, the set of constructions v-constructed by $\lambda c[B^*_{wt}\ {}^0a[{}^0Sub\ c\ {}^0d\ {}^0C(d)]]$ is non-empty, and the conclusion $[{}^0\exists\lambda c[B^*_{wt}\ {}^0a[{}^0Sub\ c{}^0d\ {}^0C(d)]]]$ v-constructs **T** as well.[36]

[35] Duží and Jespersen [7, Section 3.3] provides a detailed exposition of *Sub* and *Tr*.

[36] This rule stems from Jespersen and Duží [7] where it is labeled 'Rule$_4$'. Rule$_4$ applies to *de dicto* attitudes, such as the above analysandum "Le Verrier believes* that Vulcan causes Mercury's perturbations'. There is a different rule for the *de re* variant, which may be phrased in either of two ways: (i) "Le Verrier believes* of Vulcan that it causes Mercury's perturbations", which uses

The quantification in the conclusion states explicitly that the set of constructions such that they occur displayed inside the context C is non-empty. The rule actually borders on triviality. This fact only speaks in its favour, for quantifying into hyperintensional contexts should, *in principle*, if not technically, be no different from quantifying into extensional contexts. The technical finesse the conclusion exhibits resides in the fact that the variable c occurs *outside* the hyperintensional context of the displayed construction C. Therefore, c occurs *free* in the Composition $[^0Sub\ c\ ^0d\ ^0C(d)]$, making it amenable to λ-binding and subsequently to existential quantification.

The above rule is too weak, however, to yield the conclusion that there is an *office* f such that Le Verrier believes that f_{wt} is a P. To validate the following inference, we need a stronger rule of quantifying-in, which will be presented below:

Le Verrier believes* that Vulcan causes Mercury's perturbations

---

There is an *office* f such that Le Verrier believes* that f_{wt} causes Mercury's perturbations

The stronger rule we need is this one[37]:

$$\frac{[^0B^*_{wt}\ ^0a\ ^0C(X/g)]}{[^0\exists'\lambda f[^0B^*_{wt}\ ^0a[^0Sub[^0Tr\,f]^0g\ ^0C(g)]]]}\ \textit{Rule B}$$

The rule, however, applies only if X is already a Trivialization of a $\iota_{\tau\omega}$-object. If the office were constructed in a more complex manner than merely by means of Trivialization, then we would not be in a position to substitute its Trivialization, for a's perspective in the premise must be reproduced in the conclusion in order to retain the hyperintensional character of a's attitude. I am presupposing that X is a Trivialization in order to keep the rule manageable. Hence, let X be a Trivialization of a $\iota_{\tau\omega}$-object. Then we can *quantify into* the hyperintensional context Trivialized by $^0C(X/g)$ and *quantify over* this $\iota_{\tau\omega}$-object. The new types are: $f, g \rightarrow_v \iota_{\tau\omega}$; $\exists'/(o(o\iota_{\tau\omega}))$.

The idea behind the proof of *Rule B* is this. Let X be the Trivialization of the $\iota_{\tau\omega}$-typed object b, i.e. X is ^0b. Then $[^0Sub[^0Trf]\ ^0g\ ^0C(g)]$ $v(b/f)$-constructs the construction C in which all the occurrences of g have been replaced by occurrences of ^0b. Thus, if $[^0B^*_{wt}\ ^0a\ ^0C(X/g)]$ v-constructs $\mathbf{T}$ then $[^0B^*_{wt}\ ^0a[^0Sub[^0Tr\,f]^0g\ ^0C(g)]]$ $v(b/f)$-constructs $\mathbf{T}$. The conclusion, $[^0\exists'\lambda f[^0B^*_{wt}\ ^0a[^0Sub[^0Tr\,f]^0g\ ^0C(g)]]]$, means that there is an office f such that a believes* that such-and-such is true of f_{wt}. It does not follow that a believes* that the same is true of individual g_{wt}, even if f and g are co-occupied, or is true of individual i, even if i occupies f.

---

the active voice and introduces an anaphor, and (ii) "Vulcan is believed by Le Verrier to cause Mercury's perturbations", which uses the passive voice in order to predicate of the occupant of Vulcan the property of being believed* by Le Verrier to be such-and-such. Hyperintensional and intensional propositional attitudes de re are studied in great detail in Duží and Jespersen [5]. Non-propositional ('objectual') attitudes like calculating are studied in Duží and Jespersen [7].

[37]This rule is a slightly restricted variant of what is called 'Rule$_3$' in Duží and Jespersen [7].

The full *proof* of *Rule B* is this. According to Definitions 12.6 and 12.7, the Composition $[{}^0Sub[{}^0Tr\ x]{}^0y{}^0C(y)]$ $v(b/x)$-constructs the construction $C({}^0b/y)$ in which the occurrences of y have been replaced by ${}^0b$. Thus the following proof-steps are truth-preserving, Att^* a hyperintensional attitude:

(1)	$[Att^*_{wt}a{}^0C({}^0b/y)]$	∅
(2)	$[{}^0{=_*}\ [{}^0Sub[{}^0Tr\ x]{}^0y\ {}^0C(y)]{}^0C({}^0b/y)]$	1, Definitions 12.6, 12.7
(3)	$[Att^*_{wt}a[{}^0Sub[{}^0Tr\ x]{}^0y\ {}^0C(y)]]$	2, Leibniz
(4)	$[\lambda x[Att^*_{wt}a[{}^0Sub[{}^0Tr\ x]{}^0y\ {}^0C(y)]]{}^0b]$	3, λ-abstraction
(5)	$\sim[{}^0Empty\ \lambda x[Att^*_{wt}\ a[{}^0Sub[{}^0Tr\ x]{}^0y\ {}^0C(y)]]]$	4, Definition 12.2 iii)
(6)	$[{}^0\exists\lambda x[Att^*_{wt}\ a[{}^0Sub[{}^0Tr\ x]{}^0y\ {}^0C(y)]]]$	5, EG

We are able now to formalize the inference stated in prose above.[38] The most significant types are as follows: V (*Vulcan*) an office of type $\iota_{\tau\omega}$; LeV (*Le Verrier*) an individual of type ι; B^* a hyperintensional attitude whose complement (second argument) is a construction of a proposition; *Sub* of type $(*_1\ *_1\ *_1 *_1)$.

$$\frac{\lambda w\lambda t[{}^0B^*_{wt}\ {}^0LeV\ {}^0[\lambda w\lambda t[{}^0P_{wt}\ {}^0V_{wt}]]]}{\lambda w\lambda t[{}^0\exists'\lambda f[{}^0B^*_{wt}\ {}^0LeV[{}^0Sub[{}^0Tr\,f]{}^0g{}^0[\lambda w\lambda t[{}^0P_{wt}g_{wt}]]]]]}$$

One admissible value of f is the office of Vulcan, regardless of the empirical fact that Vulcan is actually vacant. The Composition $[{}^0Sub[{}^0Tr\,f]{}^0g{}^0[\lambda w\lambda t[{}^0P_{wt}g_{wt}]]]$ $v(V/f)$-constructs the Closure $\lambda w\lambda t[{}^0P_{wt}\ {}^0V_{wt}]$.

Note that if, in the conclusion, Le Verrier were related to the Trivialization ${}^0[{}^0Sub\ c\ {}^0V\ {}^0[\lambda w\lambda t[{}^0P_{wt}\ {}^0V_{wt}]]]$, rather than the Composition $[{}^0Sub\ c\ {}^0V\ {}^0[\lambda w\lambda t[{}^0P_{wt}\ {}^0V_{wt}]]]$, he would find himself facing the very procedure of executing the substitution, and not just its product or result. Le Verrier would need to be not only an astronomer and mathematician, but also a logician. This is one assumption we should not be making. If we want agents to carry out acts of substitution themselves, we need to add the (empirically unrealistic) premise that an agent always executes the substitution and does so flawlessly. The above analysis relates Le Verrier instead to the result of the substitution.

Also the inference does not require Le Verrier to perform an inferential act to the effect that since he believes* that Vulcan causes Mercury's perturbations then there is an office whose occupant he believes* to have this property. Le Verrier is not availing himself of inferential knowledge. Whoever does draw the above inference learns the logical nature of Vulcan, namely that Vulcan is an office and not an individual, and also that there exists at $\langle w, t\rangle$ at least one office such that Le Verrier believes* that its occupant causes Mercury's perturbations.

[38]When *analyzing* an argument bearing on a natural-language inference, we encode in symbolic notation the propositional constructions that are the respective *meaning* of the premise(s) and the conclusion. When *proving* an argument, the first step is λ-elimination, because each step is to be *truth*-preserving (valid) and not meaning-preserving.

12.4 Conclusion

Above I demonstrated how to achieve existential quantification into one sort of hyperintensional attitude context and over intensional entities of a particular type. This achievement helps move hyperintensional logic in particular and philosophical logic and formal semantics in general one step further ahead. The investigation into the validity of quantifying into hyperintensional attitude contexts serves at least two purposes. One is to make explicit an ontological commitment that is implicit in the premise. Reflecting on quantifying-in challenges us to get clear(er) about the nature of a particular part occurring inside an agent's attitude complement. The other purpose is to field quantifying-in as a touchstone for various theories of hyperintensionality: do they, or do they not, validate quantifying-in, and if they do, how do they achieve this? TIL validates quantifying-in, and it does so in a principled manner. Quantifying-in, in TIL, does not require contextualist epicycles or other ad hoc measures, but flows forth from its semantics, which sports universal transparency.

Acknowledgements The research reported herein was funded by *FP7-PEOPLE-2013-IEF* Project No. 628170 *USHP*: *Unity of Structured Hyperpropositions*, and VŠB-Technical University of Ostrava Project No. SP2014/157: *Knowledge Modeling, Process Simulation and Design*. Versions of this paper were read as an invited lecture at Munich Centre of Mathematical Philosophy, Ludwig Maximilian University, 24 October 2013, and as an invited tutorial at the Eberhard Karls University of Tübingen, 7 February 2013. I am indebted to Marie Duží for precious comments; to Alessandro Torza for his kind invitation to contribute to this volume, and granting much-needed deadline extensions, as well as helpful comments on the penultimate draft; and to Iker for making the writing process so much longer and infinitely more pleasurable.

References

1. Bealer, G. 1982. *Quality and concept*. Oxford: Clarendon.
2. Carnap, R. 1947. *Meaning and necessity*. Chicago: Chicago University Press.
3. Crawford, S. 2008. Quantifiers and propositional attitudes: Quine revisited. *Synthese* 160: 75–96.
4. Duží, M. 2012. Extensional logic of hyperintensions. *Lecture notes in computer science,* vol. 7260, 268–290.
5. Duží, M. and B. Jespersen. 2012. Transparent quantification into hyperpropositional contexts *de re*. *Logique et Analyse* 220: 513–554.
6. Duží, M. and B. Jespersen. Manuscript. Transparent quantification into hyperpropositional contexts *de dicto*.
7. Duží, M. and B. Jespersen. 2015. Transparent quantification into hyperintensional objectual attitudes. *Synthese*192: 635–677. doi:10.1007/s11229-014-0578-z.
8. Duží, M ., B. Jespersen, and P. Materna. 2010. *Procedural semantics for hyperintensional logic: Foundations and applications of transparent intensional logic*. LEUS, vol. 17, Dordrecht/London: Springer.
9. Forbes, G. 1996. Substitutivity and the coherence of quantifying in. *Philosophical Review* 105: 337–372.
10. Forbes, G. 2013. Marcus and substitutivity. *Theoria* 78: 359–374.

11. Geach, P. 1965. A medieval discussion of intentionality. In *Logic, methodology and philosophy of science*, ed. Y. Bar-Hillel, 425–433. Amsterdam: North-Holland.
12. Hanks, P.W. 2011. Structured propositions as types. *Mind* 120: 11–52.
13. Hanks, P.W. 2013. First-person propositions. *Philosophy and Phenomenological Research* 86: 155–182.
14. Jespersen, B. 2012. Recent work on structured meaning and propositional unity. *Philosophy Compass* 7: 620–630.
15. Jespersen, B. 2015. Structured lexical concepts, property modifiers, and transparent intensional logic. *Philosophical Studies* 172: 321–345. doi:10.1007/s11098-014-0305-0.
16. Kaplan, D. 1986. Opacity. In *The philosophy of W.V. Quine*, ed. L.E. Hahn, and P.A. Schilpp, 229–289. La Salle: Open Court.
17. Marcus, R.B. 1993. *Modalities*. Oxford: Oxford University Press.
18. Materna, P. 1997. Rules of existential quantification into intensional contexts. *Studia Logica* 59: 331–343.
19. Priest, G. 2002. The hooded man. *Journal of Philosophical Logic* 31: 445–467.
20. Quine, W.V. 1956. Quantifiers and propositional attitudes. *Journal of Philosophy* 53: 177–187.
21. Sainsbury, M. 2010. Intentionality without exotica. In *New essays on singular thought*, ed. R. Jeshion, 300–318. Oxford: Oxford University Press.
22. Sleigh, R.C., Jr. 1967. On quantifying into epistemic contexts. *Noûs* 1: 23–31.
23. Smullyan, R. 1948. Modality and description. *Journal of Symbolic Logic* 13: 31–37.
24. Soames, S. 2010. *What is meaning*? Princeton: Princeton University Press.
25. Sylvan, R. 2003. The importance of nonexistent objects and of intensionality in mathematics. *Philosophia Mathematica* 11: 20–52.
26. Thomason, R. 1980. A model theory for propositional attitudes. *Linguistics and Philosophy* 4: 47–70.
27. Tichý, P. 1971. An approach to intensional analysis. *Noûs* 5: 273–297.
28. Tichý, P. 1975. What do we talk about? *Philosophy of Science* 42: 80–93.
29. Tichý, P. 1978. Two kinds of intensional logic. *Epistemologia* 1: 143–164.
30. Tichý, P. 1986. Indiscernibility of identicals. *Studia Logica* 45: 251–273.
31. Tichý, P. 1988. *The foundations of Freges logic*. Berlin/New York: de Gruyter.
32. Williamson, T. 2013. *Modal logic as metaphysics*. Oxford: Oxford University Press.
33. Zimmermann, T.E. 2006. Monotonicity in opaque verbs. *Linguistics and Philosophy* 29: 715–761.

Chapter 13
Carnap, Quine, Quantification and Ontology

Gregory Lavers

Abstract At the time of *The Logical Syntax of Language* (*Syntax*), Quine was, in his own words, a disciple of Carnap's who read this work page by page as it issued from Ina Carnap's typewriter. The present paper will show that there were serious problems with how *Syntax* dealt with ontological claims. These problems were especially pronounced when Carnap attempted to deal with higher order quantification. Carnap, at the time, viewed all talk of reference as being part of the misleading material mode of speech, and as such dismissed, rather than addressed, ontological problems. Central to the analysis in the present paper is the concept of an explication, which was seen by both Carnap and Quine as being of great philosophical importance. It will be shown that the concept of explication played a significant role in how each formulated their mature position on ontology. Both these final positions on ontology can also be seen as a evolving in reaction to Carnap's flawed handling of ontological matters at the time of *Syntax*. Carnap, influenced by Tarski's work on semantics, comes to believe that the concept of reference can be given an acceptable explication, and that by doing so we can see reference to abstract objects as unobjectionable. As a result, Carnap develops a position very different from the one presented in *Syntax*. Quine strongly rejected the instrumentalism of *Syntax*, and sought to give an explication of ontological questions that was language independent. This paper closes with a discussion of each's understanding of the other's position.

G. Lavers (✉)
Department of Philosophy, Concordia University, 1455 Maisonneuve Blvd. West, PR-202, Montreal, Quebec, Canada H3G 1M8
e-mail: glavers@gmail.com

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_13

13.1 Introduction

The purpose of the present paper is to provide a clear understanding of the dispute between Carnap and Quine on matters of ontology.[1] I trace the dispute back to an unstable position, on existence assumptions in logic, which Carnap held at the time of *The Logical Syntax of Language* (*Syntax* hereafter). This problematic position is especially apparent in Carnap's understanding of higher order quantification. One of the things *Syntax* explicitly sought to achieve was to show that philosophical claims tend to be, properly understood, claims about the features of some language. Quine, in reaction to this, from very early on, attempted to formulate ontological questions in a way that was both very clear, and, at the same time, *language transcendent*. I will show that Quine took himself to be giving an explication of such terms as *ontology*. His answer, of course, is that we are ontologically committed to all the ineliminable entities in the domain of quantification of our best scientific theories. I show that Carnap's position in 'Empiricism, Semantics and Ontology' (ESO hereafter) is also tied to the notion of explication. For Carnap, to answer questions about the reference of terms for abstract objects requires a two stage explication—one at the level of the object language and one at the level of metalanguage.

The case of set theory (or type theory) is of particular importance when it comes to discussing ontology. As Quine maintains, once you have set theory, all abstract objects may then be defined as sets. Carnap thought we have a considerable amount of freedom in explicating the notion of logical truth, and he himself included type theory in this category. Quine views all set theories and type theories as non-logical, because all such theories make arbitrary stipulations about which collections exist. Since he sees higher order logic as quantifying over sets, Quine thought only first-order logic ought to count as logic. Quine's reasons for opposing higher order quantification are closely related to problems Carnap was facing in *Syntax*. Carnap, there, chose to accept (non-substitutional) higher order quantification and chose to take an instrumentalist stance towards the logical portion of the language. Quine sees Carnap's instrumentalism as illegitimate, and rejects higher order quantification because he believes the ontological price is too high.

Ultimately, Quine fails to understand the significant changes in Carnap's views on ontology after Carnap was exposed to Tarskian semantics. Carnap saw semantic meta-languages as making possible the explication of notions such as *reference*. Carnap took such an explication to show how, without engaging in any metaphysics, it is possible to understand reference to abstract objects. Quine took Carnap's mature

[1]The present paper, although independent, is something of a sequel to my paper "On the Quinean-Analyticity of Mathematical Proposition" [17]. In my earlier paper I look at how the positions of Carnap and Quine on analyticity are related to their views on explication. In that paper I had to purposefully ignore their debate on ontology in order to focus on their views on analyticity. This paper is meant to do the opposite—analyticity will be considered only where it is necessary to consider it in order to understand their respective positions on ontology.

position on ontology to be a minor reformulation of his position from *Syntax*. While Quine did not understand the role of explication in Carnap's mature position on ontology, Carnap likewise, failed to understand that Quine was attempting to provide an explication that preserves the core meaning of the term 'ontology'. Carnap saw his dispute with Quine as being purely terminological, but he was clearly wrong in this. In doing so, Carnap really fails to understand Quine's goal. That said, a worry that Carnap expressed, although it may not apply to Quine, seems clearly to apply to many that have been influenced by Quine's views on ontology.

In Sect. 13.2, I discuss the manner in which ontological assumptions are handled in *Syntax*. I show that they are handled in a way that is clearly circular, and that this problem was not discovered because of certain views Carnap held which prevented him from seeing any questions about abstract ontology as even being well formulated. Section 13.3 contains a brief discussion of Carnap and Quine on the notion of an explication. Section 13.4 will discuss Quine's formulation of a language independent way of addressing ontological questions. Section 13.5 contains a discussion of Carnap's position in ESO. Section 13.6 discusses their positions on the status of set theory as logic. The final two Sects. (13.7 and 13.8) are devoted to each's understanding of the other's position on ontology.

13.2 Carnap, *Syntax* and the Formal Mode of Speech

Carnap claims, in his 'Intellectual Autobiography' [10, p. 53], he and other members of the Vienna Circle had come to reject the Wittgensteinean view that we can say nothing of the logical structure of language. Carnap was influenced by the metamathematical work of Hilbert, Tarski and Gödel, but sought to generalize metalinguistic investigations beyond those of metamathematics. Ultimately, Carnap hoped his science of logical syntax would make clear which philosophical problems were really questions about the features of a certain language. As a book introducing a new and general method of linguistic analysis, *Syntax* begins by addressing the question of what logical syntax is. §1 opens with the following lines:

> By the **logical syntax** of a language, we mean the formal theory of the linguistic forms of that language—the systematic statement of the formal rules which govern it together with the development of the consequences that follow from these rules.
>
> A theory, a rule, a definition, or the like is to be called formal when no reference is made in it either to the meaning [*Bedeutung*] of the symbols (for example, the words) or to the sense [*Sinn*] of the expressions (e.g. the sentences), but simply on the kind and serial order of the symbols from which the expression is constructed. [2]

Notice that 'formal' here is not used as we would define it today but is defined as being equivalent with 'not concerning the sense or reference of either individual words or complete sentences'. In fact, throughout *Syntax* 'formal' is identified with being unconcerned with meaning. Carnap, at the time, views talk of meaning as part of the material (*inhaltlich*) mode of speech and responsible for much philosophical confusion.[2] Sentences that mention the meaning of terms are to be translated into the formal mode of speech, which involves eliminating the concept of meaning. At several places in *Syntax*, Carnap recognizes that, concerning the languages and definitions he is outlining, he makes significant existential assumptions. However, his various strategies for dealing with these ontological worries constitute an unsatisfactory patchwork. In this section, I would like to explore the link between Carnap's dismissing all questions concerning the meaning of terms, and his various, and not very convincing, attempts to deal with ontological questions.

When Carnap defines 'analyticity' for Language II in §34d, he does so on the basis of of what he called *valuations*.[3] The set of valuations of a given type is what we would now call the domain of quantification for that type of variable. If x is a variable of type 0, then the class of valuations for it are the accented expressions $(0, 0', 0'' \ldots)$. A valuation for a standard first level predicate variable will be any *arbitrary collection* of zero level expressions. A valuation for a variable that stands for a second level predicate is an arbitrary collection of valuations of first level predicate variables, an so on for every type in the language.[4] In this way for each type of variable there is a, usually uncountable, intended domain associated with it.

Since a valuation for a numerical variables is an accented expression, that is, an actual string of symbols, it can unproblematically be called syntactic. However, for all higher types of variables, a valuation will be a class of valuations of lower type. Carnap was aware that this at least hinted at a platonistic interpretation of higher order quantification:

> Thus the definition must not be limited to the syntactic properties which are definable in S, but must refer to all syntactic properties whatsoever. But do we not by this means arrive at a Platonic absolutism of ideas, which is non-denumerable and therefore can never be exhausted by definitions, is something that subsists in itself, independent of all construction and definition? [2, §34d]

Carnap, of course, denies that the view he is defending is platonistic. The reason he gives is that we can define the set of valuations for some language S, in a stronger syntax language S_2. Of course for this to work properly S_2 must be be interpreted in a standard way and so only pushes the problem back a step. This was pointed out to Carnap, much later, in [1] to which Carnap agreed [9].[5] At the time of *Syntax*, however, Carnap simply points to the fact that analyticity for S is defined in some distinct language as all that is required to avoid any platonistic commitments. I take it few philosophers today would view this 'but the definition can be given in another language' point to successfully eliminate the worry that too strong existence

[2] In a footnote to §56 of *Word & Object*, Quine writes 'It was indeed I, if I may reminisce, who in 1934 proposed 'material mode' to him as a translation of his German.'

[3] Sections with letters affixed to the numbers were prepared for the original German edition but not included for lack of space.

[4] For simplicity I am avoiding discussing relations and functions.

[5] Despite the 1963 publication date, most of the material for the Schilpp volume on Carnap was written in the mid-1950s. This is still, of course, much later than *Syntax*.

assumptions are being made. This is especially true since existential assumption at least as strong have to be made concerning the domain of quantification for S_2.

We have just seen that Carnap's definition of 'analyticity' for Language II involves quantification over uncountable totalities. Despite interpreting his quantifiers as ranging over such uncountable collections, Carnap believed his languages to be quite innocent of ontological problems. §38a is devoted exclusively to addressing the problem of existence assumptions in logic.

> If logic is to be independent of empirical knowledge, then it must assume nothing concerning the *existence of objects*. For this reason Wittgenstein rejected the Axiom of Infinity, which asserts the existence of an infinite number of objects. And, for kindred reasons, Russell himself did not include this axiom amongst the primitive sentences of his logic. [2, §38a, original italics]

He begins the section by exploring how one would construct a logical system that makes no existence assumptions. He then notes that his Language I and II are not such systems, but appeals to the distinction between co-ordinate languages and name languages to dismiss ontological worries. Name languages pick out elements of their domain by name, whereas co-ordinate languages pick out elements of the domain in a systematic way by using numbers. I have argued elsewhere [16] and [18] that this does nothing to address the ontological problems that he seems to be worried about, and will not go into significant detail about this here. I will only point out that the domain of quantification could be identical between a name language and a coordinate language.[6] As such it is unclear what this distinction can do to ease the concerns of those who have serious worries about ontology.

From our perspective, it may seem that the Carnap of *Syntax*, when faced with ontological worries, would simply restate the principle of tolerance:

> It is not our business to set up prohibitions, but to arrive at conventions...
>
> In logic there are no morals. Everyone is at liberty to build up his own logic, i.e. his own language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments. [2, §17]

It can't both be true that logic must assume nothing concerning the existence of objects, and that in logic there are no morals. That is, if we are free to outline any language we wish, then we ought to be able to outline languages with all kinds of statements of the form '$(\exists x)Px$' among its theorems. Carnap, however, never makes this straightforward move to defuse ontological concerns. The closest he comes is at the end of §38a where he is considering the ontological implications of accepting the axiom of choice. However, as we see, he does not merely restate the principle of tolerance as a justification for existence assumptions in logic, but asserts that the mathematical portion of the language is a mere tool for the purpose of making correct descriptive claims:

[6] Since what characterizes a coordinate language is that elements of the domain are picked out in some systematic way, there is no reason why we would be limited to countable domains. We might pick out the objects in some domain systematically using real numbers or ordinals for instance.

> The $\mathfrak{S}_{\mathfrak{l}}$ [logical sentences] (and with them all sentences of mathematics) are, from the point of view of material interpretation, expedients for the purpose of operating with $\mathfrak{S}_{\mathfrak{d}}$ [Descriptive sentences]. Thus, in laying down an $\mathfrak{S}_{\mathfrak{l}}$ as a primitive sentence, only usefulness for this purpose is to be taken into consideration. [2, §38a]

Here Carnap's answer is that, sure, the axiom of choice makes certain existential assumptions, but we are unconcerned with these since we need only care about the material interpretation of descriptive sentences. But adopting this instrumentalist view concerning existence assumptions of the logico-mathematical portion of the language is to dismiss rather than address (or dissolve) ontological concerns.

When Carnap pays some attention to the ontological commitments of his languages (or definitions), as we just saw, he employs three different strategies. In the definition of analyticity for Language II, faced with quantification over uncountable totalities, Carnap simply states that the definition can be given in a distinct language. It is hard to see anything very satisfying in this response. Later, in §38a, he appeals to the distinction between name languages and co-ordinate languages to address the ontological commitments of his language systems. Here again, his response is unsuccessful, this distinction does nothing to address the worry, since the domain of entities may be identical between a name language and a co-ordinate language. Finally we saw that he dismisses ontological concerns because we need not worry about the interpretation of the logical portion of the language. We can treat it as a mere instrument for deriving descriptive claims. The reason we can ignore the interpretation of the logical vocabulary is that all sentences involving only logical expressions are either analytic or contradictory. Indeterminateness comes in only with the descriptive vocabulary. Once the descriptive vocabulary is given a material interpretation, then every sentence will be either true or false. So once we have interpreted the descriptive vocabulary, no further interpretation is required (see §62). But now we see that his instrumentalist stance towards the logical vocabulary depends on his definition of 'analytic', and we have seen that there is very little provided in terms of an argument that this definition does not involve serious ontological assumptions (at least in the case of Language II). Recall that his argument here was merely that the definition of 'analytic in S' could be given in some distinct language S_2. Presumably, Carnap would take an instrumentalist stance towards the logical vocabulary of S_2, but then we have clearly come full circle. Taken all together then, we can see that in *Syntax* Carnap has done nothing to address those who are concerned about ontological assumptions in logic.

Carnap himself points out that there may be some cause for concern about the ontological assumptions regarding his Languages I and II. His response to these concerns, however, is, as we have just seen, to sweep them under the rug. We can now see this as quite inevitable. The reason he could not deal more satisfactorily with ontological questions is that certain views he held at the time of *Syntax*, and all of which he abandons shortly afterward, prevented him from viewing any questions about existence assumptions as even well formulated.[7] In the remainder of this section I would like to examine these views that were abandoned shortly after *Syntax*.

Let us begin then with Carnap's view that we must use what he calls *syntax languages* to explore the logical features of an object language. An object language may have any vocabulary that one wishes, but Carnap held that syntax languages are languages whose sole function is to talk about certain object languages. As such, he thought, their descriptive vocabulary, if they have any at all, should be limited to what is needed to discuss what symbols appear at which places.

Carnap realized that he could define truth for logical languages. "If S_1 is a *logical language*, then, with respect to S_1, 'true' corresponds to 'analytic'." [2, §63] He did not see how it was possible to define truth for descriptive languages, because such a definition would have to be given in a syntax language that contains little or no descriptive vocabulary. Truth then, along with meaning, are relegated to the material mode of speech. We must now look at Carnap's views concerning the misleading nature of what he calls the *material mode of speech*. The material mode of speech is characterize as involving *talk of meaning* or by the use of *universal words*. To obtain a proper understanding of a sentence, if it is not a straightforwardly empirical claim, we need to translate it into the *formal mode of speech*. Remember 'formal' is taken to mean not concerned with meaning. If we have a sentence involving the concept of meaning, say, to use Carnap's own example, "Yesterday's lecture was about Babylon", we need to translate it into one that does not involve the concept of meaning. In this case, we can translate it as 'The word 'Babylon' or a synonymous expression was used in the previous lecture'. In translating into the formal mode of speech we are also supposed to eliminate universal words. A universal word is a word for a property that holds of all the entities of a certain type (that is universally true for a certain type of variable). Assuming numbers make up a logical type, the statement "five is a number" involves a universal word. It should be translated as " 'five' is a number word". Here Carnap would call the statement "five is a number" a pseudo object sentence. A pseudo object sentence is defined as a quasi-syntactic sentence of the material mode of speech—where a sentence is quasi-syntactic if it is equivalent to a statement expressible in a syntax language.

Carnap describes translatability into the formal mode as the *touchstone* of meaningfulness for all philosophical sentences. In this section we are concerned with how Carnap addresses (or avoids addressing) ontological questions, especially concerning logical objects. But if the proper understanding of a question is obtained only once our question is formulated in the formal mode, then we see we cannot

[7][4] §39 discusses which theses of *Syntax* need to be altered in light of of developments in semantics. His general outlook here is that, on the whole, the various theses in *Syntax*, including discussions of the material mode of speech and of quasi-syntactic sentences "remain valid" but ought to be "supplemented by the corresponding semantical discussions". This is an unstable position, given the material mode of speech is predicated on the elimination of the notion of reference, and the notion of quasi-syntactic depends on the obsolete notion of a syntax language.

ask questions about the existence assumptions in logic at all. We certainly can't ask if 'five' refers to a number. This would involve both a universal word and the concept of reference. When translated into the formal mode it would become " 'five' is a numerical expression". Any hint of ontological assumptions is removed. To consider just one more example, let us look at Carnap's own translation of the claim that arithmetic involves numbers and numerical properties etc.:

10a. The sentences of arithmetic *state* (or express) certain properties of numbers and certain relations between numbers.
10b. The statements of arithmetic are composed of numerical expressions and one- or many-termed numerical predicate in such and such a way. [2, §75]

Clearly Carnap saw it as an advantage of his system that it did away with ontological questions. My goal in this section was to show how various particularities about *Syntax* prevented Carnap from being in a position to give a satisfactory answer to questions about the existence assumptions in logic. The pieces are now almost all in place to make this connection. The views discussed in the previous few paragraphs were all abandoned by Carnap in his semantic phase. These include the limitation to syntax languages and the need for translatability into the formal mode of speech. Once Carnap accepts semantic metalanguages, including a full translation of the object language, he realizes that he can't dismiss certain questions for being *quasi-semantic*, since all statements would be quasi-semantic. "Jane is over five feet tall" could be translated as "A true sentence results from substituting 'Jane' for 'x' in the predicate 'x is over five feet tall' ". I would like to point out now that already in *Syntax*, Carnap realized that for logical languages the property of being quasi-syntactic is trivial. In fact he explicitly says as much concerning logical languages: "in this case, the concept 'quasi-syntactical' becomes trivial." [2, §63] The reason for this, we can now see, is that syntactic metalanguages could contain a full translation of a logical object language. Recall that the restriction on syntactic metalanguages is that they contain *no descriptive vocabulary* (beyond that needed to say which symbols appear where). They can include all the logical vocabulary one would want. Carnap dismisses such claims as "five is a number" as being a quasi-syntactic sentences of the material mode of speech, and Carnap takes it that properly understood this becomes a question about the features of a language. But this is in *exactly the same sense* in which, once semantic metalanguages are accepted, any assertion may be seen as making a claim about a language.

We saw above that Carnap employs three strategies when dealing with ontological assumptions in logic. The distinction between name and co-ordinate languages is nothing but a red herring. The other two strategies were seen to each support the other (in a clearly circular way). He defends his definition of analyticity by claiming that the ranges for the valuations of various types could be defined in a distinct metalanguage. He then goes on to maintain an instrumentalist reading of the logical sentences of a language. The justification of this is presumably that, if true, they are analytically true. But then each of these last two strategies works only if the other does. Carnap does not seem to be aware of the problems with these various strategies. The reason for this would appear to be that Carnap views all ontological questions about the logical portion of the language as really questions about the features of certain languages. However, logical sentences are quasi-syntactic *for exactly the same reason* that all sentences become quasi-semantic once the move to semantic metalanguages is made. That is, in a trivial sense that does not succeed in showing that they are really questions about language.

The concepts of 'truth' and 'meaning' were considered to be part of the misleading material mode of speech. We have just seen that Carnap thought translation into the formal mode of speech, which lacks these concepts, was required before we could properly understand what was being claimed in sentences involving these concepts. As is now well known, what Carnap called 'syntactic' at the time of *Syntax* includes much of what we would now call semantics.[8] For instance he defines the relation of consequence, analyticity, and synonymy. But there is reason to think that even for the concepts of 'truth' and 'reference', which are so strongly associated with the material mode of speech, Carnap did not see proper definition as an impossibility:

> The material mode of speech is not itself erroneous it only readily leads itself to wrong use. *But if suitable definitions and rules are laid down and systematically applied, no obscurities or contradictions arise.* Since, however, *the word-language* is too irregular and too complicated to be actually comprehended in a system of rules, one must guard against the dangers of *the material mode of speech as it is ordinarily used in the word-language* by keeping in mind the peculiar character of its sentences. [2, §81, my italics]

Although Carnap abandoned many of the specific theses of *Syntax*, the above quote is important because it represents a view that remains constant in Carnap's philosophy. In later years he might express very much the same thought with reference to his concept of *explication*. He might say: the concepts of 'truth', 'reference' and even 'existence' are, in ordinary language, imprecise to the point of inviting fruitless philosophical disagreements; explications of these notions, on the other hand, may be very fruitful and important. Let us, now, then, turn to the subject of explications.

13.3 Carnap and Quine on Explication

Quine [27, p. 41] describes himself as "very much a disciple of Carnap's for six years". Early in this period (which extends roughly from 1933 to 1939) Quine "attended [Carnap's] lectures and read his *Logische Syntax* page by page as it issued from Ina Carnap's typewriter." [27, p. 41] [9] By 1951 Quine describes Carnap's influence over him by saying "Though no one has influenced my philosophical thought more than Carnap, an issue has persisted between us for years over the

[8] See [14] for an early argument to this effect.

[9] This would, then, have included the sections of *Syntax* prepared for the original German edition but not included for lack of space.

questions of ontology and analyticity." [21] The remainder of this paper will concern principally their disagreement in the 1950s on the subject of ontology. In several works leading up to the early 1950s [5, 6] and [7] Carnap develops his account of an *explication*. The concept of explication became a central pillar of Carnap's thought, but Quine also saw the notion of a Carnapian explication as very important. Very many of Quine's works, including many of the most important ones, contain a discussion of explication (for example, *Word & Object* [22], 'Two Dogmas ...' [20], 'Epistemology Naturalized' [25], and *The Web of Belief* [30] all contain at least some discussion of explication). I have discussed in detail the relationship between Carnap and Quine's account of explications in [17]. Here I wish only to outline their views and then demonstrate the relationship between their views on explication and their views on ontology.

Carnap's account of explication begins by rejecting a certain more traditional view of the goal of analysis. On the traditional view the goal of an analysis is to come up with a clear definition of a concept that is identical to the concept under analysis. However, if identity is required, the definition can be no more clear than the notion being analyzed and therefore analysis cannot in principle yield anything fruitful. Once this condition of identity is dropped, we see that in giving an analysis we are introducing a new notion (Carnap calls this the explicatum) in place of the already understood notion (the explicandum). Beginning with the observation that the explicandum and explicatum cannot, on pain of making no progress whatsoever, be required to be identical, Carnap goes on to impose the weakest possible condition on the relationship that must hold between them. The condition is merely that the explicatum is similar enough to the explicandum that it could usefully be used as a replacement. In his *Logical Foundations of Probability*, Carnap outlines four desiderata of an explication:

1. The explicatum is to be *similar to the explicandum* in such a way that, in most cases in which the explicandum has been so far used, the explicatum can be used; however, close similarity is not required and considerable differences are permitted.
2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an *exact* form, so as to introduce the explicatum into a well-connected system of scientific concepts.
3. The explicatum is to be a *fruitful* concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).
4. The explicatum should be as *simple* as possible; this means as simple as the more important requirements (1), (2), and (3) permit. [7, §3, original italics]

The important thing to note about these is that it is only the first desideratum that mentions the explicandum, and only the loosest relation is required between the explicandum and explicatum.

Quine wholeheartedly agrees with Carnap that we cannot require the concept arrived at after an analysis to be identical with the notion we had prior to an analysis. Carnap spoke of explication as replacing an existing concept with a new one. Quine, in what amounts to the same thing, speaks of eliminating the old troublesome concept in favour of a clear counterpart.

> A similar view can be taken of every case of explication: *explication is elimination*. We have, to begin with, an expression or form of expression that is somehow troublesome. It behaves partly like a term but not enough so, or is vague in ways that bother us, or it puts kinks in a theory or encourages one or another confusion. But also it serves certain purposes that are not to be abandoned. Then we find a way of accomplishing those same purposes through other channels, using other less troublesome forms of expression. The old perplexities are resolved. [22, §53, original italics]

Notice, however, and this is very important, in Quine's account of explication we are preserving certain features of the explicandum. The above quote is from §53 of *Word and Object*. This section is given the title 'The ordered pair as a philosophical paradigm'. Quine's point is that the various definitions of the ordered pair disagree on many points, and are in fact mutually inconsistent, but what they disagree on can be labeled 'don't cares'. More importantly, what they agree on, and what is core to their meaning, can be summed up in the following condition:

$\langle x, y\rangle = \langle w, z\rangle$ only if x=w and y=z

Concerning a proposed set theoretic definition of the ordered pair, Quine states:

> This construction is paradigmatic of what we are most typically up to when in a philosophical spirit we offer an "analysis" or "explication" of some hitherto inadequately formulated "idea" or expression. We do not claim synonymy. We do not claim to make clear and explicit what users of the language had in mind all along. We do not expose hidden meanings, as the words 'analysis' and 'explication' would suggest; we supply lacks. *We fix on the particular functions of the unclear expression that make it worth troubling about*, and then devise a substitute, clear and couched in terms of our liking, that fills those functions beyond those conditions of partial agreement, dictated by our interests and purposes, any traits of the explicans come under the head of "don't cares". [22, §53, my italics]

This is a more traditional account of explication than Carnap's. For Quine explications begin by identifying what it is about the explicandum that we wish to preserve. Only then do we provide a replacement that preserves these features. That this was *not* a feature of Carnap's conception can easily be seen in Carnap's paper 'Quine on Analyticity'.[10] Carnap wrote this paper in response to Quine's 'Two Dogmas ... ', but it was unpublished until its inclusion in [29]. Quine, in his attacks on Carnap's definitions of 'analyticity', is often looking for what features of the explicandum are preserved by the explicatum. Carnap repeatedly accuses Quine of *confusing* properties of the explicatum with those of the explicandum. For Carnap, no particular features need to be preserved. The phase of identifying the core meaning of an expression, which then needs to be preserved, is simply absent from Carnap's account. This difference in their accounts of explication is subtle, and subtle enough that neither of them seemed to notice that they did not share the same view. We will see below that understanding this difference in their views is important for understanding their respective positions on questions of ontology.

[10] Carnap does talk of explication as a two stage process. We begin by clarification of the explicandum, and then we provide the explicatum. But in the second stage we are in no way bound by what is identified in the first stage.

13.4 Quine and Ontology

Quine in his work with Goodman [15] famously tried to defend a nominalism about abstract entities. Quine, also famously, eventually came to view the nominalist project as hopeless. In this sense Quine's views on ontology certainly evolved. However, as to how to address questions of ontology, Quine's views are remarkably stable. In his 'A logistical approach to the ontological problems' [19], Quine wishes to distinguish between terms that genuinely name entities and *syncategorematic* expressions which do not.[11] The key, Quine urges, is to look at what expressions may be replaced with a variable that can then be quantified over.[12] "It thus appears suitable to describe *names* simply as those constant expressions which replace variables and are replaced by variables according to the usual laws of quantification. [...] To *be* is to be the value of a variable" [19, 199] Of course, so far, this distinction between names and syncategorematic expressions will be highly dependent on the specific features of the language with which one is dealing. However, Quine clearly wants to push further than this and arrive at something that is not purely linguistic:

> Shift of language ordinarily involves a shift of ontology. There is one important sense, however, in which the ontological question transcends linguistic convention: How economical an ontology *can* we achieve and still have a language adequate to all the purposes of science? In this form the question of the ontological presuppositions of science survives. [19, p. 201]

Remember, in *Syntax*, Carnap classified all questions about what the logical vocabulary referred to as, properly understood, questions about the features of certain languages. Quine is here searching for a way in which ontological questions are not merely questions about the features of a particular language. He believes he has arrived at a *language transcendent* manner to pose ontological questions. If, for the purpose of an adequate formulation of our scientific theories, we need to quantify over certain kinds of objects, then the claim that such things exist is not a mere feature of a particular language.

As mentioned above, the approach to ontological questions first presented in this 1939 article did not change much throughout Quine's career. We are ontologically

---

[11] Preprints of this paper were made available, and the paper was to be included in volume 9 of *Erkenntnis*, but the journal ceased publication before volume 9 was produced.

[12] [13] is an interesting review of this work of Quine's. This work hints at the nominalist project, and Church already sees its demise. Church writes "Apparently it is hoped that an adequate formalized language may be devised in which all abstract nouns are syncategorematic, and the *tenability* of the nominalistic position thereby demonstrated.

It would seem, however, that such a demonstration of the tenability of the nominalistic position must be at the same time a demonstration of its extreme artificiality. In the opinion of the reviewer, the effect is only to emphasize the illusory character of the question whether abstract nouns *really* have designata. For the matter is relative, on the present showing, not only to the choice of a particular language, but also the choice as to which particular notation or notations in the language shall be regarded as denoting the existential quantification (the syntax of the language will ordinarily not determine the latter choice uniquely)."

committed to all those entities in the domain of quantification of our best scientific theories, where ontological economy is but one of many norms within science. So for Quine, ontological questions, even when they concern logico-mathematical entities, are on par with other questions in science. We can now ask if Quine thought of this as an explication of the term 'ontology'. Of course 1939 predates Carnap's earliest discussions of explication by 6 years. But what of Quine's remarks about ontology after he was exposed to Carnap's concept of an explication? There is clear evidence that Quine did consider this to be an explication of the term 'ontology':

> Now my ethics of terminology demand, on occasion, the avoidance of a word for given purposes when the word has been pre-empted by in a prior meaning; meaningless words, however, are precisely the words I feel freest to specify meanings for. But actually my adoption of the word 'ontology' for the purpose described is not as arbitrary as I make it sound. Though no champion of traditional metaphysics, *I suspect that that the sense in which I use the word has been nuclear to its meaning all along*. [21, pp. 203–4, my italics]

When Quine says that he suspects he has identified the sense that was nuclear to the metaphysicians' use of the term, *given his views on explication*, he is stating that he suspects that he has successfully explicated the metaphysicians' use of the term 'ontology'.[13] Remember, for Quine, giving an explication consists in identifying the core (or nuclear) meaning of an expression—the part of it's traditional meaning that is clear and useful—and then giving a precise definition that preserves this feature. So in 1951 there seems to be clear evidence that Quine suspects himself to have successfully explicated the term 'ontology'.

In his 1966 paper 'Existence and Quantification', Quine begins by discussing the case of singular existence claims such as 'Socrates exists'. A traditional logical analysis of language might insist that such claims are meaningless because it is impossible to assert of an object that it exists. Quine argues that we should regard, '(∃ x)(x = Socrates)' as an explication of what we mean to express when we claim that Socrates exists. He then turns his attention to statement of the form 'Ps exists' where P is a predicate. Here we are asking about the role of the existential quantifier in statements of the form (∃ x)*P*x. Quine holds that there is no unified answer that could serve as an explication of all such cases:

> We found an explication of "a exists" as "(∃ x)(x = a)"; but explication in turn of the existential quantifier itself, "there is," "there are," explication of general existence is a forlorn cause. Further understanding we may still seek even here, but not in the form of explication. We may still ask what counts as evidence for existential quantification. [24]

This may seem to conflict with the claim above that Quine saw himself as successfully explicating the term 'ontology'. In fact, however, there is no conflict at all. The 1966 view is perfectly consistent with his 1951 view that he takes himself to have identified the core meaning of the term 'ontology'. What our theories *say* exists can be given a unified explication. It is this that his explication of ontology in terms

[13] Of course, Quine does not think it worthwhile to go through a detailed study of how metaphysicians have used the term to show that this is in fact the case.

of our domain of quantification achieves. But which existence claims we *should* accept is not a matter to be decided by explication. To this question, of course, Quine appeals to his naturalism and holism.

13.5 Carnap, Explication and *ESO*

The discussion above of Carnap's position in *Syntax* ended with his claim that there is nothing in principle wrong with the concepts particular to the material mode of speech, so long as they are given clear definition. It is only their use in ordinary language that is so unclear as to lead to philosophical confusion. We saw above that Carnap, at the time of *Syntax*, could define truth for logical languages. In such cases 'true' and 'analytic' coincide. But given his self-imposed restriction to syntactic meta-languages, he could not define truth for descriptive languages. In his 'Intellectual Autobiography' Carnap recounts the meeting where Tarski first told him of his definition of truth. Carnap says that he assumed Tarski meant logical truth, but was surprised to hear that Tarski meant our ordinary notion of truth, including truth as it applies to contingent factual claims. Carnap immediately challenges Tarski to give the truth conditions for a simple claim like "this table is black". Of course, Tarski replies "The sentence 'This table is black' is true if and only if this table is black." Carnap continues:

> In his treatise Tarski developed a general method for constructing exact definitions of truth for deductive language systems, that is, for stating rules which determine for every sentence of such a system a necessary and sufficient condition for its truth. In order to formulate these rules, it is necessary to use a metalanguage which contains the sentences of the object language or translations of them and which, therefore, may contain descriptive constants, e.g., the word "black" in the example mentioned. In this respect, the semantical metalanguages go beyond the limits of syntactical metalanguages. This new metalanguage evoked my strongest interest. I recognized that it provided for the first time the means for *precisely explicating* many concepts used in our philosophical discussions. [10, p. 60–61, my italics]

There are a couple of things to notice about this quote. First, Carnap clearly identifies the liberalization from syntactic metalanguages to semantic metalanguages as making possible the definition of truth. Secondly, and more important for our purposes, Carnap speaks of explicating many further notions used in philosophical discussions. Carnap clearly sees Tarski's definition of truth as an explication. In fact, besides Frege's definition of number, it is Carnap's most used example of a successful explication.[14] But Carnap, thinks that a definition of truth is only one

[14]Concerning Frege's explication of number Carnap writes "Before Frege, nobody was able to give an exact account of the meanings of [arithmetical] words in non-arithmetical terms. By Frege's *explication* of the numerical words, which I regard as one of the greatest philosophical achievements of the last century, the logical connection between these words and logical particles like "there is", "not", "or", and "the same as" became completely clear for the first time. Therefore we have to say that in spite of practical skill in usage, people in general, and even mathematicians

important notion that the liberalization to semantic meta-languages permits. Carnap quickly realized that semantic languages permit the definition of ‘reference’ (or ‘designates’). We simply require of a ‘designate’ predicate that all statements of the form “ ‘a’ designates a ” be provable.[15] Of course, talk of reference was the hallmark of the material mode of speech, but now we see Carnap realized even this concept is capable of clear explication.

Carnap’s 1939 ‘Foundations of Logic and Mathematics’ shows how quickly Carnap abandoned his *Syntax* thesis that we ought say nothing concerning the meaning of symbols. In §14 of this work, Carnap defines zero, the successor function, and the property of being a finite cardinal number in the manner that “Frege has shown”.[16] In §17 Carnap introduces the Peano axioms with ‘b’, ‘′’ and ‘N’ as primitives. He then goes on to say:

> The *customary interpretation* of the Peano system may first be formulated this way: ‘b’ designates the cardinal number 0; if ‘. . . ’ designates the cardinal number n then ‘. . .′’ designates the next one, i.e., $n + 1$; ‘N’ designates the class of finite cardinal numbers. Hence on this interpretation the system concerns the progression of finite cardinal numbers ordered according to magnitude. [3, 182]

In 1934, Carnap dismisses all questions about the reference of terms, but now in 1939 he is happy to talk of the terms of Peano arithmetic designating finite cardinal numbers.

The goal of the present section is to talk about the role of *explication* in ESO, but ESO has not even yet been mentioned. Although it might not seem like it, what has been discussed already is essential for introducing a discussion of ESO. What has been said so far might strike some, however, as having little to do with what transpires in ESO. For instance, the concepts of *external questions* of *linguistic frameworks* have not at all been mentioned (until just now). And in fact, I will discuss these concepts as little as possible. It is true, much of the discussion of ESO concerns these concepts. In turn, much of the secondary literature complains that these concepts are ill defined. The concepts of *external questions* and *linguistic frameworks*, were used by Carnap as a way of illustrating his mature position on matters of ontology, but they are not necessary for understanding that position. Neither were they used by Carnap outside of ESO, except when discussing the position of that paper. What is central to his mature philosophical position on ontology is his notion of explication and his view that semantic metalanguages can be used to give explications of notions (such as reference) that he previously dismissed.

---

before Frege, were not completely clear about the meaning of numerical words.” [11, p. 935, my italics]

[15]If the metalanguage does not contain the object language, but contains a translation of the object language, this condition must be adjusted accordingly.

[16]There are differences, however, between Carnap and Frege’s definition. Carnap defines the numbers as classes of the second level. It is also worth noting that Carnap is now aware that the definitions depend on a standard interpretation of the higher level quantifiers.

The term *explication* does not appear even once in ESO, but that does not mean it does not play a very important role in the paper. In fact, I would say that the paper, properly understood, is all about explications. ESO is five years after the first explicit discussion of explication, and from the same year as [7] which contains Carnap's most detailed discussion of explication. Also, 1950 is just prior to when all of the material for the Schilpp volume was prepared, and here it is clear that the notion of explication is central to how he approaches most philosophical problems. By leaving out the concept of explication from one's understanding of ESO, and focusing on the concepts of external questions and linguistic frameworks, it is hard to see this work fitting in naturally with Carnap's other writings.

I want to claim, in fact, that ESO is concerned with explications from start to finish. For instance when Carnap considers, in ESO, how we introduce the system of numbers, he talks of defining the individual numbers, the general property of being a finite number, etc. Clearly what Carnap has in mind here is a Frege-type definition of number. And, as was mentioned above, Carnap views Frege's definition of number as an exemplar of explication. Carnap is explicit about this in many places. So by talk of introducing the *framework of arithmetic*, Carnap clearly has in mind *giving an explication of our arithmetical vocabulary*—that is, providing a particular systematic treatment of number. When Carnap talks of *the system of propositions*, he has in mind an account of propositions similar to that given in *Meaning & Necessity*. He even, in a footnote to the section of ESO dealing with propositions, tries to clarify a point about this previous discussion of propositions from *Meaning & Necessity*, where Carnap clearly thinks of himself as giving an explication of the concept of proposition. "The greatest difficulty in the task of explicating the concept of proposition is involved in the case of the false proposition." [6, p. 29] The same could be said of all of the various *linguistic frameworks* that Carnap discusses. What he has in mind in each case is a formalized language that serves as an explication of a certain range of vocabulary (whether vocabulary concerning things, numbers, propositions, properties, etc.).

Most importantly, explication again comes into play with what Carnap himself identifies as the main task of the paper. Carnap says the following in the introductory section:

> Recently the problem of abstract entities has arisen again in connection with semantics, the theory of meaning and truth. Some semanticists say that certain expressions designate certain entities, and among these designated entities they include not only concrete material things but also abstract entities, e.g., properties designated by predicates and propositions designated by sentences. Others object strongly to this procedure as violating the basic principles of empiricism and leading back to a metaphysical ontology of the Platonic kind. *It is the purpose of this article to clarify this controversial issue.* [6, p. 206, my italics]

Remember, in *Syntax*, giving an interpretation of a language involves providing a material interpretation of only the descriptive vocabulary. But here he identifies as the central goal of the paper to defend the use of abstract objects as the referents of terms in a semantic theory. This goal, however, is postponed until the final section of the paper (apart form the conclusion). Here the argument is presented with such incredible brevity that it is not surprising that most commentators on

ESO do not address it at all. The argument does not even take up the entire section but is contained in only a few lines. The majority of the section contains a tangentially related discussions of Ryle and British empiricists. But let us now look at the argument —filling in the required reasoning. Carnap begins by considering a semantic claim where an abstract object stands as the referent of a term:

(a) ‘five’ designates a number.

Before we can discuss (a) we need the ‘framework of numbers’ in which both the individual numbers and the general concept of number are defined. Of course, Carnap has in mind here a Fregean definition of number, which Carnap sees as an explication of our arithmetical vocabulary. In such a system it will hold that:

(b) Five is a number.

But this Frege-style explication of our arithmetical vocabulary is not expressive enough for us to yet formulate (a). We need a semantic metalanguage for our language of arithmetic. So we introduce a metalanguage that contains a full translation of the object language. This metalanguage will include explications of our semantic vocabulary as they apply to statements of the object language:

> Further, to make the statement (a) possible, L [a meta-language for the language of arithmetic] must contain an expression like “designates” or “is a name of” for the semantic relation of designation. If suitable rules are laid down, the following is likewise analytic: (c) ‘five’ designates five. [6, p. 217]

Carnap then points out that from (b)—which results from a explication of our arithmetical vocabulary—and (c)—which results from an explication of semantic expressions as they relate to the object language—(a) is a trivial consequence.[17] Carnap then goes on to maintain that the same argument applies no matter what we start with as our object language. “Thus the question of the admissibility of entities of a certain type or of abstract entities in general as designata is reduced to the question of the acceptability of the linguistic framework for those entities.” [6, p. 217]

Everyone agrees that we can set up logical systems where we can give a Frege-style definition of number.[18] Some philosophers might, however, wish to regard such a system as nothing but an empty formalism. Sure, they may say, we could define numerical vocabulary in that way, but we should not see these terms as referring to anything. Carnap’s point is that we can introduce ‘refer’ (or ‘designate’) in the

[17] Quine, as is well known, makes a lot out of Carnap’s use of ‘analytic’ in the above quote. But Carnap could have equally used the term ‘provable’ here instead of ‘analytic’.

[18] Of course, one might say say that because of the need for a standard interpretation of higher order logic, one cannot be sure to have *completely unambiguously* defined the numbers. But whatever one’s views on higher order quantification, one cannot deny that, at least with impressive clarity, we can define such a system.

precise sense explicated in the semantic metalanguage, and it will be a theorem of such a formalized language that numerical terms refer. The view that we ought to see such terms as without reference is now seen as unmotivated. No longer is " 'five' refers to a number" hopelessly unclear metaphysics, it is now a theorem of a well defined formalized language.[19] In this way, Carnap hopes to help empirically minded philosophers "to overcome their nominalistic scruples." [6, p. 206]

I said I would discuss the concepts of *linguistic frameworks* and *external questions* as little as possible. I have already mentioned that what Carnap has in mind, when he talks of linguistic framework, is an explication of a certain range of vocabulary. For example the framework of numbers is an explication of our arithmetical vocabulary and the framework of propositions would consist of an explication of the concept 'proposition'. Let me now close with a brief remark about what an external question is. We just saw that, relative to an explication of number, numbers exist. And relative, to an explication of the notion of reference for an arithmetical object language, numerical terms refer. One might say at this point, yes, relative to this newly introduced sense of 'refer', numerical terms refer, but is this the correct sense of refer—is there actually an object for which these terms stand? Since Carnap has offered an *explication* of the term 'refers' as it relates to the object language, he would say there is no question of whether the account of reference is correct. An external question then is one that asks, of 'exists' or 'refers' in some reconstructed system, if they agree with reference and existence in the unreconstructed sense—a sense Carnap saw, in *Syntax* and right through his semantic period, as being sufficiently unclear as to invite philosophical confusion.[20]

13.6 Carnap, Quine, and Set Theory

For Carnap there are no deep mysteries in the philosophy of mathematics. Today, questions about the existence of numbers, or of whether numerical terms refer, are seen by many philosophers as quite mysterious. For Carnap, to answer such questions involves no mystery, but simply a two-stage explication. We begin by giving an explication of our arithmetical vocabulary—in a type theoretic background, for instance. We then explicate our semantic vocabulary relative to this object language. Once this is done 'Numbers exist' and 'Numerical terms refer' become theorems of the appropriate formalized language.

[19] According to an explication of our arithmetical vocabulary, and an explication of our semantic vocabulary as it applies to our system of arithmetic, numbers exist and numerical terms refer. Carnap does not take this position to amount to *platonism*. Platonism would involve asserting that numbers exist and numerical terms refer, in an unexplicated sense of 'exist' and 'refers' (technically, in giving an explication of our arithmetical vocabulary we do not explicate existence, but show the connection between logical notions like existential quantification and our arithmetical vocabulary—see Footnote 14).

[20] Howard Stein briefly makes a similar point about external questions being questions concerning the correctness of an explication (see [32] p. 280).

Of course, as mentioned, these explications have to take place in a background theory—be it type theory (with an axiom of infinity) or set theory, or something else.[21] What can we say about the status of this background theory? Carnap saw Frege, Hilbert, and Russell and Whitehead, for instance, as all involved in the project of of explicating the notion of logical truth. Carnap himself tended to prefer type theoretic languages, and explicitly states that the notion of L-truth he defines relative to these languages is meant as an explication of the notion of logical truth. So the status of the background theory is that it is itself an explication of our concept of logical truth. The question of whether all of type theory is really part of logic, is a question about the correctness of such an explication, and given Carnap's conception of an explication, it is not a legitimate question.

Notice how much turns on the explication of 'logic'. Carnap's account of explication requires only *similarity* between the explicatum and the explicandum. Type theory, complete with higher order quantification, is certainly similar to what has traditionally been called logic, and so Carnap intends to count it as such. Carnap, therefore, views the project of logicism a having already been successfully carried out by Frege. All that was needed was to import Frege's work into a consistent background theory. Quine, of course, does not count set theory (or type theory) as part of logic. We will turn shortly to the question of why Quine did not see set theory as part of logic. First however, I should say something now about Carnap's preference for type theoretic languages. In a letter to Quine, Carnap explains his preference for (often many-sorted) type-theoretic languages:

> I feel somewhat uneasy when entities like Socrates, kindness, & 7 are grouped together as "objects". Frege did so, and it was his undoing. You can, of course, avoid contradictions by suitable restrictions. But the question is whether the contradictions are not symptoms for a fundamental unsoundness. [29, 1947-4-13]

Interestingly, Quine responds to this this very point in saying:

> I agree that the logical antinomies are symptoms of a fundamental unsoundness somewhere, but I suspect that this unsoundness lies in platonism itself—i.e., in the admission of abstract values of bindable variables. The contradictions which issue from platonism can indeed be staved off by various artificial devices, and in my view the theory of types is merely one such artificial device. [29, 1947-5-1]

Carnap sees us as skirting inconsistency by grouping too many intuitively distinct kinds of objects into one all encompassing domain. We will see later, that Quine took Carnap's preference for such languages to be based on his desire to preserve his prejudice against *universal words*. Quine held this position for many years, even through the 1960s, but this suspicion on Quine's part is without merit. In *Meaning & Necessity* Carnap explicitly rejects a prejudice against universal words as unwarranted.

---

[21]Carnap discusses the axiom of infinity in §37e of [8]. Here he says that it can either be taken as a primitive sentence—an axiom, or taken as a rule in the meta-language that makes the assertion of the existence of infinitely many objects L-true. It is clear from here (and from [10] pp 47–48) that Carnap never had a definitive position on the axiom of infinity, but thought that under the proper interpretation it should count as analytic.

Quine's response, just quoted, to Carnap's preference for type theories leads nicely into our discussion of why Quine rejected any kind of higher order quantification as part of logic. In the quote above, Quine expresses worries about quantification over abstract entities, and also expresses his belief that type theory is merely an "artificial device". Quine's rejection of second-order logic *as logic* is tied to his views on set theory.[22] Quine, in many places in his writing, expresses the same argument against set theory (or type theory). The argument is that we have one intuitive notion of set and that is the notion of set introduced by naïve comprehension. The paradoxes show this notion of set to be inconsistent, and all further developments of set theories or type theories are simply ad hoc devices designed to avoid paradox. That is to say, various set theories and type theory are not an explication of our intuitive notion of set, since they do not preserve the defining feature of our intuitive notion of set (naïve comprehension). Consider for example:

> But we cannot simply withhold each antinomy-producing membership condition and assume classes corresponding to the rest. The trouble is that there are membership conditions corresponding to each of which, by itself, we can innocuously assume a class, and yet these together yield a contradiction. We are driven to seeking optimum consistent combinations of existence assumptions, and consequently there is a great variety of proposals for the foundations of general set theory. Each proposal is unnatural, because the natural scheme is the unrestricted one that the antinomies discredit; and each has advantages, in power and simplicity or in attractive consequences in special directions, that its rivals lack. [28, p. 16]

In §55 of *Word & Object* Quine begins by saying that if we have sets, then we have all we could ever need, because any other abstract object could be explicated in set theory. He then goes on to give the same argument that there's only one natural comprehension principle and many ad hoc ones. But, so far, these are arguments against a set theory (or type theory, since he sees this too as an ad hoc means of avoiding paradox) in general, and not an argument as to why they do not count as logic. In 'Carnap and Logical Truth' Quine sketches how the argument concerning the ad hoc nature of set theory can be extended to an argument that set theory is not part of logic:

> I will not here review the important contrast between logic and set theory, except for the following one. Every truth of elementary logic is obvious (whatever this really means), or can be made so by a series of individually obvious steps. Set theory, in its present state anyway, is otherwise. [. . . N]o consistent set theory is both adequate to the purposes envisioned for set theory and capable of substantiation by steps of obvious reasoning from obviously true principles. What we do is develop one or another set theory by obvious reasoning, or elementary logic, from unobvious first principles which are set down, whether for good or the time being, by something very like convention. [23, p. 388]

So here we get one answer as to why set theory might not count as logic. Quine takes it as a feature of our intuitive notion of logic that it must involve reasoning by obvious steps from obvious (in some sense) first principles, and then shows that, whatever we mean by obvious, set theory fails this test. Of course, Quine is not

[22] For further discussion of Quine's views on set theory and higher order logic see [31].

putting forward, as a serious theory, that logic proceeds from obvious steps from obvious first principles. His main aim is to show that Carnap's 'linguistic doctrine' of logical truth is no more an explanation of how we know logical truths than the view that logic is obvious. For this reason Quine does not go into detail about what he means by 'obvious'. But despite the not fully worked out nature of the account, this argument does give us insight into why Quine thought set theory was not logic. Set theory is not logic because it proceeds from non-obvious (arbitrarily stipulated) conventions. But given that these reasons for not including set theory as logic are based on a sketch of a criterion, which Carnap points out [12], as it stands, does not even rule out 'I have five fingers on my hand' as a logical truth, it can hardly be seen as a definitive argument.

There is another argument, in his *Philosophy of Logic*, for why set theory (and higher order logic) are not properly parts of logic. Here Quine defines logical truth as a truth such that sentences with the same grammatical structure is also true. That is to say a true sentence is a *logical truth* if truth is preserved over any substitution on its atomic components. Quine shows, for first order languages, assuming the language is expressive enough, this definition coincides with other definitions of logical truth such as being true in all models. He then argues that because set theoretic truths and truths of higher order logic cannot be captured substitutionally, they ought not be considered logical truths. Higher order quantifiers must be seen as either quantifying over attributes (intensions) or over sets (extensions). Quine clearly sees ontological economy as a norm for logic. Logic should make minimal ontological demands even at the level of metatheory. It is for this reason that he proposes to capture logical truth substitutionally instead of talking about models. It is also for this reason that he rejects the 'staggering existential assumptions' of set theory and higher order quantification.

In this work, Quine is dealing with the same issues that Carnap faced in *Syntax*. There Carnap thought logic should make minimal existence assumptions, and had originally wanted to interpret higher order quantification substitutionally. Gödel, however, showed him that it would not work. Carnap's answer was to accept higher order quantification as quantification over uncountably many arbitrary sets, but to, at the same time, take an instrumentalist stance toward these existence claims. We saw that there are serious problems with the way ontological claims are dealt with in *Syntax*, and we also saw how Carnap's position on these matters changed in response to the development of semantics. Quine, as we will see in the next section, continued to see Carnap as holding a version of the *Syntax* position on existential assumption in logic. It is for this reason that Quine sees Carnap as helping himself to existence assumptions without being willing to pay the ontological price.

Quine's substitutional understanding of logical truth did not become standard, but his view that set theory (or type theory) is not part of logic did become standard (and largely due to his influence). Up until the 1950s most systems of logic did assume sets, extensions or other similar notions. So an explication of logical truth that includes such a notion is not a break from historical precedent. This is not to argue for a return to the view that set theory is logic, but merely to demonstrate that, at the time, it would not have seemed as unnatural as it does today to claim that logic includes set theory or type theory.

13.7 Quine's Understanding of Carnap on Ontology

If we were to describe Carnap's *Syntax* period views on questions of ontology in just two principles, one would be the necessity for translation into the formal mode of speech, the other would be the need to take only an instrumentalist stance towards the logical vocabulary. The translation into the formal mode of speech involved the elimination of universal words. The instrumental stance toward the logical portion of the language was supported by Carnap's position that all logical sentences are either analytic or contradictory and so not in need of material interpretation.

When Quine discusses Carnap's mature views on ontology he sees them as a mere minor reformulation of his earlier views. Consider, for example his discussion in 'Ontological Relativity'. "In his later writing this doctrine of universal words takes the form of a distinction between internal and external questions, in which people come to grips with the relative merits of theories." [26, p. 52] Quine goes on to attack the earlier view by saying that universal words are identified by their meaning ('number' is a universal word, but the extensionally equivalent predicate 'less than seven or greater than five' is thought to be unproblematic). Given his views on meaning, Quine doubts that such a distinction can be made. He then, without discussing the matter in more detail, proclaims that the 'internal' / 'external' distinction fares no better.

In his 'Carnap's views on ontology', Quine also makes it clear that he sees Carnap's internal/external question distinction as reformulation of the *Syntax* position on universal words.

> But now I want to examine the dichotomy which, as we see, underlies Carnap's distinction between external and internal, and which I am phrasing as the distinction between category questions and subclass questions. It is evident that the question whether there are numbers will be a category question only with respect to languages which appropriate a separate style of variables for the exclusive purpose of referring to numbers. [21, 207–208]

To rephrase 'external questions' as 'category questions' is to assume that what is wrong external questions is their use of universal words words. This quote is from 1951 and the one considered just before was from 'Ontological Relativity' which was originally presented in 1968. So Quine believed for at least 18 years that the position of ESO was a fairly minor modification of the *Syntax* position on existential assumptions in logic. It is important, however, to note something else about this last quote. We also see here reference to Carnap's preference for type theoretic languages through the talk of separate styles of variables. Quine continues:

> [Carnap] is thinking of languages which contain fundamentally segregated styles of variables before any definitional abbreviations; and he is thinking of styles of variables that are sealed off from one another so utterly that it is commonly ungrammatical to use a variable of one style where a variable of another style would be grammatical. A language which exploits this sort of basic compartmentalization of variables is that of Russell's theory of types. However, I think many of us overstress the theory of types to the neglect of its coeval alternative, Zermelo's set theory and its descendants.

Now, it is true that Carnap did prefer type-theoretic languages. But this attitude of Carnap's, that distinct kinds of things should be assigned distinct logical types has, of course, nothing to do with his former views concerning universal words. In fact, in response to Quine's comments on an early draft of *Meaning & Necessity*, Carnap adds in the published version:

> It is important to emphasize the point just made that, once you admit certain variables, you are bound to admit the corresponding universal concept. It seems some philosophers (not Quine) overlook this fact; they do not hesitate to admit into the language of science variables of the customary kinds, like sentence variables ('*p*', '*q*', etc.), numerical variables, perhaps also predicate variables of at least level one, and other kinds; at the same time, however, they feel strong misgivings against words like 'proposition', 'number', 'property' (or 'class'), 'function', etc. because they suspect in these words the dangers of an absolutist metaphysics.[23] [6, p.44]

Here, Carnap is clearly agreeing with Quine that we can formulate, for any given type, a universal predicate for that type. That is, Carnap is here stating that his previous position with regards to universal words is untenable—for any type, there is a definable universal predicate for that type and thus no reason to have any prejudice against terms like 'number', 'property' etc.

We have seen that Quine interprets the mature Carnap as trying to maintain some version of his *Syntax* position against universal words. We began this section by saying that the position in *Syntax* on ontology had two main components. First is the necessity of translation into the formal mode of speech—including the elimination of universal words. The second is the instumentalist stance towards the logical portion of the language. This, as we saw, was supported by Carnap's view that because sentences of the logical portion of the language are analytic (or contradictory) no interpretation of this portion of the language is required. When Quine relates the rejection of the concept of analyticity to considerations of ontology, he takes this to block a certain move on Carnap's part. Quine takes it that Carnap wants to divide existential claims into two groups which Quine calls *emprirical* and *ontological* existence claims, in order to then ignore the ontological existence claims on the ground that they are analytic. Consider:

> The contrast that [Carnap] wants between those ontological statements and empirical existence statements such as 'there are black swans' is clinched by the distinction between analytic and synthetic. [21, p. 210]

or again:

> Carnap [...] has recognized that he is able to preserve a double standard for ontological questions and scientific hypotheses only by assuming an absolute distinction between the analytic and the synthetic; and I need not say again that this is a distinction which I reject.

[23] That Quine, more than 20 years after the publication of *Meaning & Necessity*, still took Carnap to be defending a version of his thesis that philosophical confusion results from the use of universal words, is reason to suspect Quine never reread the published version to see how Carnap responded to his comments on the early draft.

> The issue over there being classes seems more a question of convenient conceptual scheme; the issue over there being centaurs, or brick houses on Elm Street, seems more a question of fact. But I have been urging that this difference is only one of degree[.] [20, pp. 45–46]

Quine understands Carnap as needing the analytic/synthetic distinction in order to make a division in types of existence claims so that he may ultimately dismiss questions about abstract ontology. Again, then, Quine is taking Carnap's mature position on ontology to be essentially the same as the position in *Syntax*. In *Syntax* Carnap has a clear double standard towards existence claims. He recognizes that he is making existential assumptions in the logical portion of the language, but as we saw, employs several strategies to dismiss these assumptions rather than address them. On the other hand the descriptive portion stands in need of a material interpretation. By the time of ESO, Carnap does not need a way to avoid dealing with existential assumptions concerning abstract objects. Given an explication of, for instance, our arithmetical vocabulary and given an explication of our semantic notions relative to that systematic account of number, the statement that numbers exist and that numerical terms refer become theorems of the appropriate formalized languages. It is true Carnap takes claims about abstract objects to be analytic. Of course, Carnap and Quine had very different views on the epistemology of mathematics and the empirical sciences, and analyticity played an important epistemlogical role for Carnap. But the concept of analyticity was not meant to support taking a dismissive stance towards all analytic existence claims. That was a view Carnap held at the time of *Syntax*, but it was abandoned shortly after. As we saw, Carnap quickly comes to see the instrumentalist position he defended in *Syntax* as unmotivated.

Carnap maintained that to use, for instance, the language of set theory is a practical decision of language choice. Quine interprets this to mean that talk of sets is a mere manner of speaking. Of course Quine did not think that Carnap was entitled to this position if it could not be shown that quantification over sets was eliminable from our best scientific theories. But Carnap did not think talk of sets was *a mere manner of speaking*. To do so would be to hold that we prove that many sets exist while working in some system of set theory, *and also hold* that sets do not exist according to the ordinary notion of existence in natural language. But Carnap takes no position on whether sets exist in the ordinary sense of existence, because he takes this notion to be unclear. There is nothing *mere* about the existence of sets for Carnap.

Quine's arguments, even in the 1960s, against Carnap's views on ontology are all, in reality, directed toward the position of *Syntax*. Quine, it seems, never recognized the (double) role of explication in Carnap's mature views on ontology. This is too bad, since Quine thought of explication as a very useful philosophical/scientific activity. As it stands, Quine thought there was something clearly illegitimate about Carnap's position on ontology. This is due to his reading the position of *Syntax* into Carnap's later works. I am not claiming that had Quine understood the role of explication in Carnap's later views he would have agreed with them, but I am trying to provide a better understanding of where their true differences lie.

13.8 Carnap's Understanding of Quine on Ontology

We have just seen that Quine seemed not to have realized the role played by explication in Carnap's mature views on ontology. It can also be said that Carnap did not realize the role played by explication in Quine's views on ontology. Carnap often suggested that his differences with what Quine says about ontology are purely terminological. Carnap accepts Quine's position that to be is to be the value of a quantified variable, but dislikes the way Quine relates this position to traditional *ontological* debates over *nominalism* and *realism*:

> I, like many other empiricists, regard the alleged questions and answers occurring in the traditional realism-nominalism controversy, concerning the ontological reality of universals or any other kind of entity, as psuedo-questions and pseudo-statements devoid of any cognitive meaning. I agree, of course, with Quine that the problem of "Nominalism" as he interprets it is a meaningful problem it is the question of whether all of natural science can be expressed in a "nominalistic" language, that is, one containing only individual variables whose values are concrete objects, not classes, properties, and the like. However, I am doubtful whether it is advisable to transfer to this new problem in logic or semantics the label 'nominalism' which stems from an old metaphysical problem. [6, p. 43]

However, it is not simply the case that Quine is giving new acceptable meanings to terms like 'nominalism' or 'ontology' from metaphysics. We saw as early as 1939, Quine is seeking a *language transcendent* way of asking about the existence of an entity. In 1951 he writes, speaking of the word 'ontology', "I suspect that that the sense in which I use the word has been nuclear to its meaning all along." [21, pp. 204] Given Quine takes an explication to involve identifying a core use that is to be preserved, this is a clear statement that Quine thought of himself as having explicated the term 'ontology'. Of course, this talk of identifying the core meaning of a term is absent from Carnap's account of explication. It is no surprise then Carnap does not understand that Quine is offering what he takes to be an explication of the term 'ontology'.

From the time of *Syntax* Carnap warns of "the dangers of the material mode of speech as it is ordinarily used in the word-language." [2, §81] That is, Carnap takes questions about the existence of objects or the reference of terms, as posed in ordinary language, to be so unclear as to invite philosophical confusion. This position is preserved in his later views. We cannot answer questions of existence and reference before explicating a certain range of vocabulary, and then explicating various semantic notions as they apply to the explication of that vocabulary. It is a basic feature of explications that they are not correct or incorrect. Since the notion of correctness does not apply, there is no further, sufficiently clear question that needs to be addressed according to Carnap. Quine's goal was to rehabilitate the very question Carnap always dismissed as a psuedo-question. The difference then, between Carnap and Quine, is clearly not merely terminological.

Furthermore, to understand how Quine intended to rehabilitate this general question of existence, we need to look again at the difference in their accounts of explication. For Carnap, once we have given our two stage (object language and metalanguage) explication and come to accept " 'five' refers to a number", there is of course no question of whether this is correct in some further sense. Explications are not to be evaluated in terms of correctness, but in terms of usefulness. On Quine's view, we begin an explication by identifying the core meaning of a term—it is then a requirement of an explication that it preserve this core meaning. Everything besides this core meaning falls under the heading 'don't cares'. The explicandum and exclicatum, of course, are not required to be identical, but they do, for Quine, need to agree on the core meaning. When Quine says he suspects that he has identified the nuclear meaning of the term 'ontology', this amount to his claiming that he has identified what *any* explication of 'ontology' ought to preserve. Any explication of what we take to exist must view us as committed to all those entities we ineliminably quantify over in our best scientific theories. This is not just one explication among many, as it would be on Carnap's account, but a general requirement on any explication of our ontological commitments.

Carnap failed to fully understand Quine's position and took their differences to be terminological. Quine was amazingly ingenious in his attempts to rehabilitate the general question of existence that Carnap dismissed. Quine saw something important preserved by his use of the terms 'ontology' and 'nominalism'. Quine was not trying to identify exactly what metaphysicians meant by these terms, but does think he has identified a core meaning that is useful and preserved by his use of the terms. Consider Quine's formulation of the problem of nominalism. Can we reformulate all of science in a language that does not involve quantification over abstract entities? Carnap, as we saw, agrees that this is a meaningful question, but sees any connection to the old problem of nominalism as undesirable. Quine is unhappy with language specific answer to existence claims—language A quantifies over abstract objects, but language B does not—and seeks a language independent way of posing ontological questions. The reformulation of the question of nominalism is a case in point. By asking if there is any nominalistic language suitable for the purposes of science, Quine has severed the ties between this problem of nominalisnm and any specific language.

Despite all of Quine's ingenuity in trying to rehabilitate a language transcendent way to address ontological questions and despite Carnap mistakenly taking their differences to be mainly terminological, a certain worry we saw Carnap expressing above is entirely justified. Carnap was worried that by using the existing term 'nominalism' for the program Quine describes, many might view an answer to Quine's question of nominalism as an answer to the traditional question of nominalism. That is, we are likely to draw a stonger conclusion than we are really entitled to. Consider the sentence 'nominalism is false if quantification over abstract objects is not in principle eliminable from our best (ideal) scientific theories'. We don't learn *something else* about the world when we learn that the notion of set is ineliminable from our best scientific theories. Given that Quine has *eliminated* the old notion of nominalism in favour of a clear counterpart, the sentence we were considering is equivalent to 'Quantification over abstract objects is not in principle eliminable from our best (ideal) scientific theories if quantification over abstract objects is not in principle eliminable from our best (ideal) scientific theories.' When Carnap says, as just quoted, "I am doubtful whether it is advisable to transfer to this new problem in logic or semantics the label 'nominalism' ... ", he is expressing the worry that one might view an answer to Quine's clearly expressed problem as an answer to the old unclear problem—even if, as Quine thinks, there is something preserved between the two, they are not identical. We must forget about all features of the old notion that are not part of Quine's explication—after all, remember, explication is elimination. I am not claiming that Quine is under any illusions about this, but certainly many people influenced by Quine take it that we would learn *something else* about the world if we were to learn that real numbers, for instance, are ineliminable from our best scientific theories.

13.9 Conclusions

One of the goals of *Syntax* was to show which philosophical questions were really questions about the features of a certain language. All questions about the logical portion of the language were labeled quasi-syntactic, and so all questions about the abstract ontology assumed by the language of science are ill-posed. By the time of ESO, Carnap thought that an explication of 'reference' could be given. It could be shown that, relative to this explication, there was no motivation for the nominalistic scruples held by many empiricists. Carnap *did not* attempt to show that talk of numbers, sets or propositions was a *mere manner of speaking*. His goal was to show how we can speak in very clear terms about abstract objects as the referents of terms. Quine understood Carnap as continuing to hold a position on ontology similar to the one at the time of *Syntax*. Quine wanted to reformulate ontological questions so as to be independent of any particular language. Carnap was worried that some might view an answer to Quine's reformulation of, for instance, the question of nominalism as an answer to the metaphysical question of nominalism. Quine, held that *explication is elimination*, and so was himself unlikely to fall prey to what Carnap was worried about. Carnap accepted that Quine had formulated a problem that is independent of the features of any specific language, but thought that making the connection to the traditional problem of nominalism might lead some to think that *something more* had been established. In the intervening years since this dispute unfolded, a Quinean approach to questions of ontology has become quite standard, and it is difficult to see, from our perspective, Carnap's worries as unwarranted.

Acknowledgements "I would like to thank Eliot Michaelson for reading a draft of this paper and offering comments. I would also like to thank Richard Creath for helping on one particular point."

References

1. Beth, E.W. 1963. Carnap's views of the advantage of constructed language systems over natural language in the philosophy of science. In *The philosophy of Rudolf Carnap, vol. XI of the library of living philosophers*, ed. P.A. Schilpp, 469–502. LaSalle: Open Court.
2. Carnap, R. 1934/1937. *The logical syntax of language*. London: Routledge & Kegan Paul.

3. Carnap, R. 1939/1955. Foundations of logic and mathematics. In *International encyclopaedia of unified science*, ed. O. Neurath, R. Carnap, and C. Morris. Chicago: University of Chicago Press, combined ed.
4. Carnap, R. 1942. *Introduction to semantics*. Cambridge: Harvard University Press.
5. Carnap, R. 1945. The two concepts of probability: The problem of probability. *Philosophy and Phenomenological Research* 5(4): 513–532.
6. Carnap, R. 1947/1956. *Meaning and necessity: A study in semantics and modal logic*, 2nd ed. Chicago/London: University of Chicago Press.
7. Carnap, R. 1950. *Logical foundations of probability*. Chicago: University of Chicago Press.
8. Carnap, R. 1958. *Introduction to symbolic logic and its applications*. New York: Dover. Revised english translation ed.
9. Carnap, R. 1963a. E. W. Beth on constructed language systems. In *The philosophy of Rudolf Carnap, vol. XI of library of living philosophers*, ed. P.A. Schilpp, 927–932. La Salle: Open Court.
10. Carnap, R. 1963b. Intellectual autobiography. In *The philosophy of Rudolf Carnap, vol. XI of library of living philosophers*, ed. P.A. Schilpp, 927–932. La Salle: Open Court.
11. Carnap, R. 1963c. P. F. Strawson on linguistic naturalism. In *The philosophy of Rudolf Carnap, vol. XI of library of living philosophers*, ed. P.A. Schilpp, 933–940. La Salle: Open Court.
12. Carnap, R. 1963d. W.V. Quine on logical truth. In textitThe philosophy of Rudolf Carnap, vol. XI of library of living philosophers, ed. P.A. Schilpp, 915–922. La Salle: Open Court.
13. Church, A. 1939. Review of W.V.O. Quine's 'A logistical approach to the ontological problem'. *The Journal of Symbolic Logic* 4(4): 170.
14. Creath, R. 1990. The unimportance of semantics. In *Proceedings of the 1990 biennial meeting of the philosophy of science association*, East Lansing, vol. 2, ed. M.F.A. Fine and L. Wessels.
15. Goodman, N., and W.V. Quine. 1947. Steps toward a constructive nominalism. *The Journal of Symbolic Logic* 12(4): 105–22.
16. Lavers, G. 2004. Carnap, semantics and ontology. *Erkenntnis* 60(3): 295–316.
17. Lavers, G. 2012. On the Quinean-analyticity of mathematical propositions. *Philosophical Studies* 159(2): 299–319.
18. Lavers, G. 2015. *Carnap on abstract and theoretical entities*. Ontology after Carnap. Prepared for: ed. S. Blatti and S. Lapointe. Oxford: Oxford University Press.
19. Quine, W.V.O. 1939/1976. A logistical approach to the ontological problem. In *The ways of paradox and other essays*, 107–32. Cambridge: Harvard University Press.
20. Quine, W.V.O. 1951/1963. Two dogmas of empiricism. In *From a logical point of view*, 20–46. New York: Harper & Row.
21. Quine, W.V.O. 1951/1976. Carnap's views on ontology. In *The ways of paradox and other essays*, 203–211. Cambridge: Harvard University Press.
22. Quine, W.V.O. 1960. *Word and object*. Cambridge: MIT.
23. Quine, W.V.O. 1963. Carnap and logical truth. In *The philosophy of Rudolf Carnap, vol. xi of library of living philosophers*, ed. P.A. Schilpp, 385–406. La Salle: Open Court.
24. Quine, W.V.O. 1966/1969. Existence and quantification. In *Ontological relativity and other essays*, 91–113. New York: Columbia University Press.
25. Quine, W.V.O. 1969a. Epistemology naturalized. In *Ontological relativity and other essays*, 69–90. New York: Columbia University Press.
26. Quine, W.V.O. 1969b. Ontological relativity. In *Ontological relativity and other essays*, 26–68. New York: Columbia University Press.
27. Quine, W.V.O. 1970/1976. Homage to Carnap. In *The ways of paradox and other essays*, 40–43. Cambridge: Harvard University Press. Revised and enlarged edition ed.
28. Quine, W.V.O. 1976. The ways of paradox. In *The ways of paradox and other essays*, 1–18. Cambridge: Harvard University Press.

29. Quine, W.V.O., and R. Carnap. 1990. Dear Carnap, Dear Van: The quine-carnap correspondence. Berkeley: University of California Press.
30. Quine, W.V.O., and J.S. Ullian. 1970/1978. The web of belief, 2nd ed. New York: Random House.
31. Shapiro, S. 1991. *Foundations without foundationalism*. Oxford: Oxford University Press.
32. Stein, H. 1992. Was Carnap entirely wrong after all? *Synthese* 93(1–2): 275–95.

Chapter 14
Quantifier Variance, Intensionality, and Metaphysical Merit

David Liebesman

Abstract Attempting to deflate ontological debates, the proponent of Quantifier Variance (QV) claims that there are multiple quantifier meanings of equal metaphysical merit. According to Hirsch—the main proponent of QV—metaphysical merit should be understood intensionally: two languages have equal merit if they allow us to express the same possibilities. I examine the notion of metaphysical merit and its purported link to intensionality. That link, I argue, should not be supported by adopting an intensional theory of semantic content. Rather, I give a general strategy for supporting claims about metaphysical merit and examine whether that strategy can be used to link merit and intensionality. Though I don't deliver a definitive verdict, the discussion provides a clearer framework for articulating and evaluating claims about metaphysical merit.

14.1 Metaphysical Merit

Myriad journal pages have been dedicated to debating whether tables exist. Such disputes strike some as a waste of time and energy, and apt for deflation.[1] Sider identifies what he takes to be the "go-to move" for a deflationist:

1. The deflationist observes a certain metaphysical dispute, in which one of the contested views is expressed by a certain sentence S.

[1]There are a number of ways such debates could be misguided. To make just one distinction, the questions asked themselves could be misguided, or the sorts of considerations countenanced could be misguided (even if the questions themselves are good). I'll focus only on deflation via the go-to move.

D. Liebesman (✉)
Department of Philosophy, The University of Calgary, Social Sciences Building, Room 1202, Calgary, AB, T2N 1N4, Canada
e-mail: david.liebesman@ucalgary.ca

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_14

2. He argues that there is an interpretation of the language of S—a way of assigning meaning to the sentences of that language—under which everyone can agree that S is true.
3. And he argues for a certain parity between this and rival interpretations [5, p. 67–68].

Sider emphasizes the importance of parity. If there were reason to favor one of these interpretations over the other, as far as metaphysics is concerned, then there would be reason to take the dispute to be settled in favor of whichever position is true on the favored interpretation.[2]

The most influential recent deflationary position, Quantifier Variance (QV), utilizes the go-to move. Consider a disputed sentence in a debate about whether tables exist:

(S) There are tables.

According to the proponent of QV, there is an interpretation of S on which it is true iff there are simples arranged tablewise. Assuming that the existence of simples is common ground, the believer in tables and the table-skeptic agree that, on this interpretation, S is true. Furthermore, according to the proponent of QV, there is nothing to favor this interpretation over myriad others. The existence of these interpretations is supported by claiming that there are multiple quantifier meanings:

(QV) There are multiple distinct meanings for the quantifiers, of equal metaphysical merit.

In this articulation of QV, equal metaphysical merit plays the role of parity in the go-to move. The proponent of QV and, more generally, anybody who pursues the go-to move, must substantiate the notion of metaphysical merit or parity in a way that supports deflationism. Those who resist deflationism will likewise want a notion of parity or merit that supports the failure of the go-to move. The importance of parity has been stressed both by Sider [5, p. 69], our paradigmatic anti-deflationist, and Hirsch [3, p. xv], the most forceful proponent of QV.

In opposing deflationism, Sider [4, 5] develops a distinctive conception of metaphysical merit. His conception has two components. The first is a commitment to the metaphysics of structure. Crudely put, the idea is that some expressions carve at nature's joints better than others. Sider's notion of joint-carving is connected to a number of issues in philosophy, and these connections purport to endow the notion with substance. An example will help. Intuitively, the predicate "blue" captures a similarity while "bleen" doesn't. (Something is bleen iff it is blue and observed before a certain time, or green and not so-observed.) The second is a commitment to the claim that, as far as metaphysical inquiry is concerned, it is a virtue for a theory to employ joint-carving terms. From these two commitments it follows that a language's metaphysical merit is determined (at least partly) by the joint-carving status of its terms.

[2] This reason may not be decisive, but it would complicate the go-to move.

Hirsch's primary target is the second commitment. Even if there are facts about structure of the sort that Sider invokes, Hirsch denies that there is any 'purely metaphysical' sense in which a language that employs joint-carving is superior.[3]

In order to substantiate this claim, the proponent of QV must provide us with an alternative to Sider's notion of metaphysical merit.[4] Hirsch construes metaphysical merit in intensional terms. In particular, he claims that any two languages with the same intensional resources are of equal metaphysical merit.[5]

Intensional resources are determined by possibility-expressing power. Take a set of possible worlds to provide the truth-conditions for a sentence. Two languages differ in intensional resources iff there is some set of possible worlds that provides the truth conditions for a sentence of one language, but no sentence in the other. If two languages have the same possibility-expressing power they are intensionally equivalent.

A question remains: why should we understand metaphysical merit in intensional terms? One answer, familiar from the literature, immediately suggests itself: we tie merit to content and adopt an intensional view of content.

I'll address two primary questions in this paper: should the proponent of QV adopt an intensional view of content, and, if not, how should the proponent of QV understand metaphysical merit? My answer to the first is no. My answer to the second consists of outlining a strategy—the equivalence-class strategy—that can be used to develop a deflationist-friendly account of merit. Particular utilizations of this strategy lead to particular deflationary conclusions. I'll focus on one utilization of the strategy with the potential to vindicate the link between merit and intensionality and outline some challenges it faces. While I won't reach a definitive verdict on the link between merit and intensionality, I will provide a framework that allows clearer articulation and more tractable evaluation of claims about metaphysical merit.

14.2 Intensional Content

On an intensional view of content, necessarily equivalent sentences express the same proposition. (Here and in what follows, I'll ignore the effects of context.) The most familiar version of this view identifies propositions with sets of possible worlds. Insofar as one wishes to link merit to intensionality, adopting an intensional view of content is tempting. In this section, I'll argue that the temptation should be resisted.

[3]This is an upshot of Chs. 3 and 4 of Hirsch [2]. He reiterated the view recently [3, xiii].

[4]Another option for the proponent of QV is to accept Sider's account of merit, but deny that any of the quantifier meanings carves at the joints better than the others. This is the gloss that Sider gives of QV in his recent discussion [5, p. 175]. Since Hirsch explicitly rejects this gloss, Sider's arguments target a different position.

[5]In fact, in his most recent discussion, he claims that QV can be derived from this assumption along with a denial of necessity invariantism: the view that there is only one possible quantifier meaning.

Why would an intensional view of content support the view that intensionally equivalent languages have equal merit? Here is an argument from the former to the latter:

1. Assume an intensional view of content.
2. From (1), intensionally equivalent languages can express all and only the same propositions/truths/contents.
3. The metaphysical merit of a language is determined by which propositions/truths/contents it allows us to express.
4. Therefore, intensionally equivalent languages have equal metaphysical merit.

Importantly, note that Sider rejects premise (3). He claims that merit is determined by how well terms carve at the joints, and this can differ among intensionally equivalent languages. So, insofar as the proponent of QV wishes to provide a genuine alternative to Sider's account of metaphysical merit, it looks as if adopting an intensional view of content doesn't help; using an intensional view of content to support an intensional view of merit requires assuming that Sider's view is false, in the form of affirming (3).

Even setting that aside, the proponent of QV should not adopt an intensional view of content. In the next two subsections I'll give two arguments for this claim. The first is that QV requires distinguishing restricted from unrestricted quantifiers, and this distinction is hyperintensional. The second is that on the most familiar intensional view of content, proponents of QV will overgenerate deflationary conclusions. While neither of these arguments is decisive, they should compel us to explore alternative ways to construe metaphysical merit.

Before getting to the details, there's a general reason to think that QV doesn't cohere particularly well with an intensional theory of content. I take the arguments to be two ways of making this general reason precise. The reason is that QV trades on the idea that there are a multiplicity of meanings that correspond to a single state of the world; which of these meanings we choose is irrelevant to capturing the state. Such a view most straightforwardly combines with views on which meaning is fairly fine-grained. This way we are assured of a multiplicity of meanings. An intensional view of content, however, is fairly coarse-grained. It eliminates distinctions made by other meaning theories. Given the paucity of meanings on an intensional view, it shouldn't be surprising that it doesn't combine comfortably with QV.

This is impressionistic. However, insofar as one finds the impression compelling, it can provide a more general explanation of the sorts of problems that derive from combining QV with an intensional view of content.

14.2.1 Hyperintensionality and the Articulation of QV

The first objection to combining QV with an intensional view of content is driven by two main claims (This section is inspired by section 9.5.1 of Sider [5]. However, the emphasis and arguments differ.):

(AQV) Articulating QV requires distinguishing between restricted and unrestricted quantifiers.

(HI) The distinction between restricted and unrestricted quantifiers requires hyperintensional content.

The upshot of AQV and HI is that articulating QV requires a hyperintensional view of content.

AQV derives some initial plausibility from the fact that Hirsch repeatedly asserts it.[6] In numerous passages, he emphasizes that the variance he takes to be important is variance of unrestricted quantifiers. We can see why Hirsch would hold this view by considering what would happen if QV merely stated that there are a multiplicity of quantifiers—restricted or unrestricted—of equal metaphysical merit. So construed, QV will not only be uncontroversially true, it will also be toothless.

I can tacitly restrict the universal quantifier so that it only ranges over my groceries ("Everything is in the fridge"). You can do the same. Thus we have two different restricted quantifiers. Furthermore, given that it is no better or worse for the purposes of metaphysics to range over my groceries than yours, these quantifiers are of equal merit. So, allowing restricted quantifiers, QV is very easy to vindicate. Too easy, in fact, because QV, so vindicated, is toothless in deflating metaphysical debates. The mere fact that we could use various restricted quantifiers that are equally good (or bad) for the purposes of metaphysics, doesn't show that there is anything defective about a debate using our unrestricted quantifier. The upshot is that we should follow Hirsch and take "quantifier" in QV to mean unrestricted quantifier.

According to HI, distinguishing between restricted and unrestricted quantifiers, which we've now shown is required in order to articulate QV, requires hyperintensional content. To motivate HI, I'll first flesh out our intensional view of content.

We've already seen that on an intensional view of content, necessarily equivalent sentences express the same proposition. We can extend this view of content to subsentential expressions in the natural way by claiming that two expressions have the same content iff they can be intersubstituted preserving content of the sentences in which they occur. The idea is that, on an intensional view of content, we can see the meanings of words as functions from sentential contexts to propositions, construed intensionally. To anticipate an objection, one could claim that this is a particularly strong intensional view of content. I'll return to its strength.

Now, we can make a case for HI by showing that there is a pair of quantifiers that, by our definition of intensional content, have the same intensional content and yet differ in whether they are restricted or unrestricted. Begin by considering the familiar metaphysical debate between common-sense ontologists (CS) and compositional nihilists (CN). Assume that both agree that there are simples, and that they agree about the basic properties of the simples and the relations between them. The proponent of QV provides us with two languages: the language in which

[6] See Hirsch [3, pp. 107, 136].

the utterances of CS theorists are true, call it CSL, and the language in which the utterances of CN theorists are true, call it CNL. Subscripting to make the language explicit, it is clear that "There are$_{csl}$" and "There are$_{cnl}$" have different contents. Substituting the latter for the former takes "There are tables" from true to false. Speakers of CSL can stipulatively introduce a restricted quantifier "There are$_{cslr}$" as follows "There are$_{cslr}$" behaves exactly like the ordinary universal quantifier, except that its domain is restricted to simples. By design, "There are$_{cslr}$" is intensionally equivalent to "There are$_{cnl}$".[7] However, the former is restricted, while the latter is unrestricted. Thus, if we are going to be able to adequately distinguish between restricted and unrestricted quantifiers, we need a hyperintensional view of content.

This completes the case for the claim that hyperintensional content is required for the articulation of QV. However, there are a number of salient worries worth addressing.

The first worry is that a more subtle intensional view of content may be able to overcome the objection. The view would combine an intensional view of sentential content, on which necessarily equivalent sentences express the same propositions, with a hyperintensional view of sub-sentential content on which we can distinguish the content of words that can be intersubstituted preserving sentential content. If this view is adopted, we can distinguish the content of "There are$_{cslr}$" and "There are$_{cnl}$" even if we admit that the sentences in which they occur express all and only the same position.

A dialectical problem with this more subtle intensional view is that it would need to be independently motivated. It is not motivation enough merely that it allows us to combine QV with an intensional view of content; we'd also want QV-independent motivation in order to adopt the view.

Another problem is that the view robs us of the resources to understand why it was important to distinguish restricted and unrestricted quantification in the first place. I argued above that we need to adopt this distinction to avoid trivializing QV. If word meaning and sentence meaning are either both intensional or both hyperintensional, we can explain what lead to the trivialization. The explanation is that while there are myriad restricted quantifiers of equal merit, they all limit our expressive ability. However, if we sever word meaning and sentence meaning, there will be no reason to think that a restricted quantifier meaning limits our expressive ability. This is brought out by "There are$_{cslr}$" and "There are$_{cnl}$": by hypothesis, they have different meanings but they allow expression of all and only the same thoughts.

Both of these problems are somewhat dialectical in nature, and they are far from decisive. However, there is a more general point worth appreciating. We're considering adoption of an intensional view of content as a method for defending an intensional view of merit. So, the goal is not merely to show that QV could, in principle, combine with an intensional view. Rather, it is to show that the combina-

[7] A few extra stipulations may have to be made to ensure that "There are$_{cslr}$" and "There are$_{cnl}$" are intersubstitutable preserving truth in all contexts. However, none of these will undermine the plausibility of claiming that "There are$_{cslr}$" is restricted.

tion is plausible and attractive. The fact that some sub-sentential hyperintensionality must be recognized makes combining an intensional view of sentential content and QV look ad hoc. This, by itself, is enough to warrant exploring other ways to make sense of metaphysical merit.

The second worry is that we should reject the straightforward view that "There are$_{cslr}$" is restricted and "There are$_{cnl}$" is unrestricted. If we relativize restriction to languages, then we can say that both quantifiers are restricted relative to CSL and neither quantifier is restricted relative to CNL. This allows us to recognize the importance of restriction without divorcing "There are$_{cslr}$" and "There are$_{cnl}$".

Two major challenges arise for a relativized view of restricted quantification. The first challenge is that it is revisionary: we usually make sense of restriction in terms of domains but the relativized approach rejects that usual stance. Thus, it must provide us with an alternate. Until the details are fleshed out, it is hard to evaluate the relativized view.

The second challenge is that, on the relativized view, there are difficulties in articulating QV. According to our articulation of QV, there are multiple unrestricted quantifier meanings of equal metaphysical merit. The problem is that if we understand restriction as relativized to languages, the result is that, assuming we are speaking English, QV is equivalent to the thesis there are multiple unrestricted quantifier meanings, relative to English, of equal metaphysical merit. By hypothesis, "There are$_{cnl}$" is restricted relative to English. So, its existence doesn't vindicate QV. In fact, it is hard to see how the sorts of quantifiers that the proponent of QV usually uses to deflate ontological disputes will be relevant to it at all, given the relativized view of quantifier restriction. Perhaps there is a way to recast QV that is more friendly to the relativistic view, but I'm unaware of it.

The third worry for the argument is that the objection obscures the fact that proponents of intensional views of content have various sophisticated strategies for dealing with what appears to be hyperintensionality. Perhaps, the worry proceeds, HI can be undermined by adopting such a strategy.

In absence of the details it is very hard to evaluate this third worry. However, I see no reason for prima facie optimism that the strategies available will be able to explain away the apparent hyperintensionality. Furthermore, I think there is substantial reason for pessimism: the best-developed strategy for dealing with apparent hyperintensionality on an intensional view does not cohere with QV. I'll now turn to this.

14.2.2 Reinterpretation Strategies and Deflationism

The second reason that a proponent of QV should be wary of adopting an intensional theory of content is that the best-developed intensional view of content overgenerates deflationary conclusions.

To understand this, it will help to look at the go-to move in some more detail. The go-to move contained three steps. Step one: identify a disputed sentence. Step two: identify an interpretation of that sentence on which disputants agree that it is true. Step three: argue for a parity between this and other interpretations.

The reason that the go-to move deflates a dispute is that the dispute can be settled solely by determining which of many candidate meanings happens to be operative, and that this metalinguistic determination itself has no metaphysical importance given the parity of the interpretations.

The result is that if we endorse the go-to move, then any debate that has the following two features will be deflated: (1) it can be settled wholly by attending to metalinguistic facts, and (2) this determination is metaphysically unimportant. I'll now argue that, on the best-developed view of intensional content, myriad debates satisfy (1) and (2), at least by the deflationist's lights.

A familiar problem with an intensional view of content is that it makes obscure what is at issue in debates that concern necessary truth or falsehood. Consider a debate about an abstruse mathematical sentence that is either necessarily true or necessarily false. Given that the disputants agree that $2 + 2 = 4$ and agree in rejecting $2 + 2 = 5$, they already converge on the content of the abstruse sentence. It follows that reasonable dispute about the sentence cannot consist in dispute about its content which, after all, is accepted by all parties.

In order to make sense of such disputes, the proponent of intensional content claims that the debate concerning the abstruse sentence does not concern its semantic content. Rather, the debate concerns some other (intensionally individuated) proposition. The process of moving from the semantic content of the sentence to the proposition being debated is reinterpretation.

Stalnaker [6, 7] has developed the most important theory of reinterpretation. His theory has the virtue of being connected to independent issues concerning the nature of assertion. Crucially, Stalnaker's reinterpretation strategy is metalinguistic. Take our abstruse mathematical sentence M. On the intensional view, the disputants are not reasonably debating the semantic content of M. However, they are ignorant of a metalinguistic fact: just which proposition is expressed by M. For all the disputants know, M expresses a truth, and for all they know, it expresses a falsehood. This yields a natural reinterpretation: take the disputed proposition to be a function from a world to the truth-value of the sentence in that world. Given that M expresses a truth in some worlds and a falsehood in others, the reinterpreted proposition is contingent. I have moved through these details very quickly, given that Stalnaker's strategy is familiar and better-presented elsewhere. However, the crucial observation is that the reinterpretation yields the result that the dispute, in some sense, concerns linguistic facts rather than the semantic content of M itself.

Now reconsider the dispute between the proponents of CN and CS. At first glance, it may seem as if Stalnaker's interpretation strategy compliments QV perfectly. Assuming that the facts about composition are necessary, the proponent of intensional content will have to reinterpret. Utilizing Stalnaker's strategy, a term about which the disputants are ignorant must be identified. The proponent of QV has a natural candidate: "There is". Given that the proponent of QV takes there to be multiple equally good meanings for "There is", they can claim that the dispute is settled merely by attending to metalinguistic facts, which are of no metaphysical importance. Hence, they will claim, the dispute is deflated.

On second glance, accepting Stalnaker's reinterpretation strategy overcommits the proponent of QV. If the proponent of QV claims that the need for metalinguistic reinterpretation deflates the debate over composition, they will claim that it deflates any other dispute in which it is needed. Consider, again, a debate over a mathematical sentence M. We've already seen that Stalnaker takes the debate over M to be metalinguistic: this is guaranteed by his reinterpretation strategy. So, the debate concerning M satisfies feature (1).

What about feature (2)? To show that the debate satisfies feature (2), we need to show that settling the metalinguistic debate isn't of metaphysical importance. The view that we're considering is driven by the claim that the metaphysical merit of a language is determined by its intension-expressing power. So, establishing that settling the metalinguistic debate doesn't affect intension-expressing power will, in this context, show that the debate satisfies (2).

One can conceive of settling a metalinguistic debate in two ways. On the first way, we settle the metalinguistic debate from an independent perspective, so to speak. In other words, we, as onlookers, determine which of many potential meanings disputants happen to be using. On the second way, we settle the debate from within: the disputants themselves determine which of many meanings they are employing. I'll now argue that neither way of settling the debate is of metaphysical importance (at least by the deflationist's lights) because neither adds intensional expressive power.

To see that the first manner of settling the debate satisfies feature (2), we need to invoke one specific feature of Stalnaker's account of assertion: an assertion must express the same proposition relative to any world in the context set. The context set is the set of worlds that the speakers take to be open possibilities; worlds that they haven't ruled out as being actual. In other words, relative to any world in the context set, the same proposition is expressed by M. It follows that which meaning for M we happen to be using doesn't affect our intensional expressive power: we express the same proposition no matter which world we occupy.

To see that the second manner of settling the debate doesn't enhance intensional expressive power, assume that discovering that we are in W allows us to express intension I with sentence M (rather than some reinterpretation). Note that it is very easy to express I even without discovering we are in world W: we could suppose that we are in W. Supposing we are in world W allows us to express the same intensions that we would be able to express by discovering that we are in world W.

Either way, settling the metalinguistic debate over the meaning of M doesn't enhance our intensional expressive power. Since the adoption of intensional content was driven by the deflationist's inclination that merit is determined by intensional expressive power, the debate satisfies (2).

Note that this exact same argument can be given in any case in which there is reinterpretation, as long as the reinterpretation proceeds in Stalnaker's manner. The ultimate result is that the go-to move sits uneasily with an intensional view of content that is developed in Stalnaker's manner: combining the two forces us to adopt a deflationary view towards any debate that involves reinterpretation.

Of course, one could pursue alternate reinterpretation strategies. There are well-known worries about Stalnaker's strategy and one may take the upshot of this section to be that we can add one more reason to the many for rejecting Stalnaker's strategy. However, even taking that to be the lesson of the argument, it is clear that the proponent of the go-to move must provide some reinterpretation strategy that doesn't conflict with the go-to move, and that this is hardly a trivial task.

14.2.3 The Upshot

Combining QV with an intensional theory of content seemed to allow us to make sense of metaphysical merit in a fairly straightforward way. Given that intensionally equivalent languages allow us to express all and only the same propositions, it was natural to take them to be equally good for the purposes of metaphysics.

However, I've now given two reasons that the proponent of QV should recognize hyperintensionality. The first is that, despite initial appearances, the notion of metaphysical merit requires hyperintensionality. The second is that the best-developed intensional view of content would overgenerate deflationary conclusions. As I already acknowledged, there may be further rejoinders one could give to these arguments. However, given that there was a very general reason to reject combining QV with an intensional view of content, and that the rejoinders would require complicated views that we don't have on the table, it is sensible to look elsewhere for a QV-friendly notion of metaphysical merit.

14.3 The Equivalence-Class Strategy

Rejecting an intensional view of content leaves us with a lacuna in understanding QV and, more generally, deflationism (at least deflationism driven by the go-to move). We must provide a QV-friendly notion of merit. Reconsider CSL and CNL, the languages in which the assertions of the common-sense ontologist and compositional nihilists are true, respectively. Given a hyperintensional view of content, these languages allow expression of different propositions. What sorts of considerations could we advance in favor of the view that they are equally suited for the purposes of metaphysics?

There is a familiar intuition that has the potential to make this problem tractable. The intuition is that the sentence "There is a table" in CSL and the sentence "There are simples arranged tablewise" in CNL seem, to the deflationist, to express the same underlying facts even if they have distinct meanings.

Here is an initial way to try and substantiate this. Distinguish between facts and propositions. For the sake of illustration we can think of propositions as modes of presentation of facts. Even if "There is a table" in CSL and "There are simples arranged tablewise" in CNL express different propositions, those propositions may present the same fact. One could then argue that the relevant notion of metaphysical merit is determined by fact-presenting capability rather than proposition-expressing capability. Even if CNL and CSL allow us to express different propositions, those propositions may present the same facts. The strategy I will now articulate for making sense of metaphysical merit departs from this idea.

14.3.1 The Strategy Outlined

Begin with the set of all propositions, and take propositions to be individuated hyperintensionally. There are numerous hyperintensional views—the most familiar distinction is between Fregean and Russellian views—though which we adopt doesn't matter for our purposes. Now partition that set. In other words, take the initial set and divide it into jointly exhaustive and mutually exclusive equivalence-classes. Two propositions are equivalent relative to a partition *P* just in case *P* sorts them into the same equivalence-class. Given this, we can define a notion of expressive equivalence relative to a partition:

> L1 and L2 are expressively equivalent relative to a partition $P =_{def}$ For all propositions *S*, if *S* is expressible in L1 then there is a proposition S' that is equivalent to *S* relative to *P* such that S' is expressible in L2, and vice-versa.

A particular thesis about metaphysical merit, then, will be tied to identifying a partition. On this view, two languages are of equal metaphysical merit just in case they are expressively equivalent relative to that partition.

This general strategy is the equivalence-class strategy, and particular executions of it—identifications of partitioning principles—yield instances of the strategy. The strategy is compatible with a hyperintensional view of content.

In the remainder of this subsection I'll draw out a number of connections between the equivalence-class strategy, the nature of facts, Sider's view of merit, and philosophical methodology. These connections will help clarify the strategy.

In building up to presenting the strategy, I invoked facts. The strategy is most natural given a distinction between facts and propositions, combined with the view that distinct propositions correspond to (perhaps by presenting) the same fact. However, the strategy does not require such a view. In fact, the strategy does not require reification of facts at all.

Assume there are no facts. We can nonetheless sort propositions into equivalence-classes. In principle, one could defend the view that a certain partition on the set of propositions is linked to metaphysical merit, despite an absence of facts corresponding to the equivalence-classes in the partition. One could even push this further and pursue the view without reifying propositions. Assuming that there are sentences, one could sort sentences into equivalence-classes with a (non-propositional) notion of synonymy, and then sort these equivalence-classes into equivalence-classes in order to understand metaphysical merit.

That said, the strategy is most naturally combined with a view that recognizes facts. The general idea would be that each equivalence-class of propositions corresponds to a single fact. This, in turn, is naturally combined with a view on which facts have multiple distinct decompositions. To understand this view, set aside facts for a moment and consider an ordinary complex material objects: my kitchen table. My table is composed of four legs and a top. There is a sense in which that decomposition is complete: it doesn't leave anything out of the constitution of my table.[8] However, there are alternative complete decompositions of the table. Consider the decomposition of my table into its left half and right half. That decomposition is a little less ordinary, but no less complete: no components of the table are left out. One could take facts to be analogous. The fact that there is a table can be multiply decomposed. One decomposition is best captured by the proposition that there is a table, and the other is best captured by the proposition that there are simples arranged tablewise.

There is nothing in the equivalence-class strategy that forces this view of facts. One could combine the strategy with a view of facts on which they don't have constituents at all, let alone multiple decompositions. More generally, we can distinguish two issues that are apt to be conflated. The first concerns the individuation of facts: how coarse or fine-grained are they? The second concerns the structure of facts: do they have constituents? In principle, one can vary answers to these questions independently and many different answers can be pursued alongside the equivalence-class strategy.

It is worth comparing Sider's view of metaphysical merit to the view employed by a deflationist utilizing the equivalence-class strategy. According to Sider, certain terms carve at the joints better than others and a language is superior insofar as it contains joint-carving terms. Even if Sider were to accept that a fact could be presented by a large number of distinct propositions, he would deny that all that matters for the purposes of metaphysics is that a language allows us to express one or another proposition that presents the fact. Rather, he would single out a particular proposition as best matching reality's structure. A prominent feature—perhaps the prominent feature—of a deflationary position based on the equivalence-class strategy is a denial that one or other ways of expressing a given fact is superior for the purposes of metaphysics. There is a crucial point here that is worth repeating. Denying that one or another language is superior for the purposes of metaphysics does not require rejecting naturalness or structure more generally. One could perfectly well accept naturalness or structure, while nonetheless denying that a language with natural terms is superior for the purposes of metaphysics. As Hirsch puts it:

[8]There are different ways we may try to make the notion of a complete decomposition more precise. The most natural uses mereological notions. For our purposes it suffices to keep it intuitive.

> Sider, the most prominent opponent of quantifier variantism, holds that (a) the world contains a natural quantification structure, and (b) there is a uniquely best ontological language. I want to define quantifier variance as the denial of (b), not as the denial of (a) [3, p. xiii]

Next, consider the methodology for settling claims about merit, given the equivalence-class strategy. I began this paper by considering what may have seemed like a very obscure question: what determines the metaphysical merit of languages? Sider's particular stand on merit—that it is determined by whether terms are joint-carving—made this question more tractable. He connects joint-carving to myriad other notions and these connections can help us evaluate his hypothesis. The equivalence-class strategy, while friendlier to the deflationist, also makes the question more tractable by linking the deflationist's position to other considerations.

Consider an execution of the equivalence-class strategy: a hypothesis that a particular partition P determines metaphysical merit. How could one support this hypothesis? Broadly speaking, there are two ways. The first is to link the hypothesis to a general theory of fact-individuation. If, for independent reasons, one argues in favor of intensionally individuated facts, then it would be natural to hold that the partitioning principle relevant to determining merit is intensional. Though issues involving fact-individuation are themselves difficult to settle, the mutual connections between fact-individuation and the equivalence-class strategy could give us a foot in the door in both cases. The second is to link the hypothesis to more specific metaphysical considerations. For instance, one may follow Fine (2000) in arguing that relational facts are identical to facts involving their converses, e.g. my being above my chair is the same fact as my chair's being beneath me. While this metaphysics of relational facts does not yield a full partition on the set of propositions, it will yield some partial results: the two aforementioned propositions will be partitioned into the same cell. This, in turn, will give us some insight into the relevant partitioning principle. In all likelihood, successfully pursuing the strategy will require both sorts of considerations. The fact that metaontological disputes become linked to other disputes is, I take it, a methodological advantage of the strategy.

Finally, it is worth remembering that the equivalence-class strategy is just that: a strategy. Particular deflationary (or inflationary) conclusions can only be supported by particular instances of the strategy. The next natural question, then, is what sort of instance of the strategy is defensible; what sort of partition on the set of hyperintensional propositions best captures metaphysical merit?

In the rest of this section I'll consider the partition that would vindicate Hirsch's claim that merit is linked to intensionality. Though I'll raise a number of key challenges, my aim is not to undermine this instance of the strategy. Rather, it is to shed light on the strategy by examining one familiar instance.

14.3.2 *Intensional Partitions*

A common theme in Hirsch's work is that unstructured facts are fundamental and an ability to express unstructured facts is what matters as far as metaphysical merit is concerned. Hirsch makes it clear that he has in mind the view that facts are intensionally individuated: necessarily co-obtaining facts are identical.[9] As I already mentioned, we can distinguish theses about structure from theses about individuation, and Hirsch's focus is the latter.

The thesis that facts are intensionally individuated generates an instance of the equivalence-class strategy with the partitioning principle being necessary equivalence. The question, then, is how this proposal fares.

14.3.2.1 Overgeneration Worries

A knee-jerk reaction to an intensional partition is that it will overgenerate deflationary conclusions. In this section I'll argue that this reaction is hasty and there are a number of ways that such a deflationist may block overgeneration.

Begin with the notion of metaphysical merit. On the envisioned version of deflationism, any two intensionally equivalent languages are of equal metaphysical merit. This is Hirsch's view, which he traces to Urmson [8]. The equivalence-class strategy allows us to flesh out and motivate this view. Facts, by hypothesis, are intensionally individuated. Propositions are sorted into equivalence-classes based on which facts they present. Two languages are of equal metaphysical merit iff they can express all and only the same facts. This is all perfectly compatible with the existence of distinct necessarily equivalent propositions.

The first sort of overgeneration worry attacks this account of metaphysical merit. Given any set of necessary propositions and any language that allows us to express more than one, we can find a language of equal metaphysical merit that expresses only one. This is counterintuitive: it may seem that a language that can express myriad necessary truths regarding universals, free will, morality, etc., is superior to the language that can only express the proposition that $2 + 2 = 4$.

There is little doubt that most theorists will share this reaction about metaphysical merit. However, the reaction may be attenuated by reflecting on the relationship between metaphysical merit and deflationism. The go-to move contained three steps. The notion of metaphysical merit only came in at the third step. In order to deflate debates about free will, universals, etc., all three steps must be established. Given this, there is room for the deflationist to insist that such debates are substantial, despite the fact that adding the ability to express additional necessary truths doesn't, in and of itself, increase the metaphysical merit of a language.

[9] See, especially, Hirsch [3, pp. 208–11].

The initial worry was that partitioning propositions by necessarily equivalence overgenerated conclusions about equivalent metaphysical merit. I've combated this by driving a wedge between conclusions about merit and deflationary results. This, in and of itself, doesn't undermine the initial reaction, which concerned metaphysical merit directly rather than the deflationary results it can be used in deriving. A skeptic may react, then, by claiming that overgeneration worries remain. I suggest that a deflationist responds to such reaction with the following sort of speech: *the notion of metaphysical merit is not one to which we have direct access. Rather, particular claims about merit are to be evaluated by their role in deflationary arguments. If a claim about merit generates the result that an obviously substantive debate is deflated, then we have grounds that it is false.*

Reactions to this speech will vary. At the end of the day, though, I suspect the deflationist will be perfectly happy to live with somewhat counterintuitive consequences regarding metaphysical merit, as long as they don't generate false deflationary conclusions. Given that, we should turn our attention to evaluating the charge that grouping propositions by necessary equivalence overgenerates such conclusions.

The second overgeneration worry is just this: that grouping propositions by necessary equivalence overgenerates deflationary conclusions. Recall, once again, the three steps in the deflationary strategy: (1) Identify a disputed sentence S, (2) provide an interpretation on which all parties agree that S is true, (3) argue for parity between this and rival interpretations. Thus far, the bulk of our discussion has focused on fleshing out the notion of parity that occurs in step (3). Let me now turn to step (2).

Step (2) requires identifying an interpretation for the disputed sentence on which all parties agree that it is true. Given the equivalence-class strategy, we can assume that an interpretation of a sentence consists of assigning it a proposition. Given that there is some proposition on which the disputants agree, there is some assignment of a proposition to the disputed sentence on which they will agree that the sentence is true. Allowing *any* assignment of propositions to count as an interpretation in the sense relevant to (2) will trivialize the step. What we need are some additional constraints. Providing constraints on interpretation is certainly not my aim in this paper, though some constraints on interpretation are independently plausible and well supported: (a) charity, (b) relative naturalness, (c) compositionality, etc. Unless one can provide an interpretation that is plausible given independently plausible metasemantic constraints, step (2) will fail to be satisfied.

Now focus on a metaphysical debate between Platonists and Nominalists. A disputed sentence in this debate is "There exists an abstract object." The envisioned deflationist will hold that, on either disputant's account, this sentence adds no metaphysical merit to the rest of ordinary English: it is either necessarily true or necessarily false. However, this will not yield a deflationary conclusion unless we can provide an interpretation for the sentence on which both parties agree that it is true. Given that the interpretation will have to be compositional and relatively natural, there is plausibly no interpretation such that the nominalist agrees that the sentence is true. To make this more vivid, consider the interpretation on which “There exists an abstract object” expresses “According to the Platonist, there exists an abstract object”. Both parties will agree that the sentence is true on this interpretation, but it is nearly impossible to see how this interpretation could be derived in a compositionally plausible manner. Since it fails to satisfy our independent constraints on interpretation, step (2) is not satisfied.[10]

The upshot is that theses about metaphysical merit, in and of themselves, don’t overgenerate deflationary conclusions. In order to argue that a particular partitioning principle overgenerates much more will have to be said. Furthermore, the deflationist can block such overgeneration by placing constraints on what counts as relevant interpretation. So, the knee-jerk reaction that partitioning by necessary equivalence will overgenerate deflationary conclusions faces serious obstacles. I’ll now turn to two other worries for such a partition.

14.3.2.2 The Individuation of Facts

As I made clear in my presentation of the equivalence-class strategy, the strategy can be pursued without commitment to facts at all, let alone a particular conception of facts. However, it is reasonable to ask for the motivation for a partitioning principle. If that principle is necessary equivalence, it is natural to provide that motivation by adopting an intensional view of facts. There are familiar problems with such a view. I’ll briefly mention three.

First, it is relatively common for facts to be identified with true propositions. This view has been motivated both by considerations of parsimony, as well as linguistic data, though the linguistic data is not unequivocal. For instance, if propositions are taken to be the meanings of sentences, it looks as if certain anaphoric pronouns that designate facts also designate propositions, given that sentences are their linguistic antecedents. For instance, “that” in the following sentence appears to designate a fact as well as a proposition: “Grass is green. That is a fact.” (Note, however, that a Fregean may hold that “Grass is green” expresses a proposition while referring to a fact, potentially diffusing this argument.) If facts are true propositions, and propositions are individuated hyperintensionally, then facts are not intensionally individuated. Given that, as I argued in Sect. 14.2, the proponent of QV should recognize hyperintensionally individuated propositions, it seems that the proponent of QV who uses the equivalence-class strategy to make sense of metaphysical merit is barred from identifying facts with true propositions.

That said, there are numerous arguments against identifying facts with true propositions. We cannot freely intersubstitute fact-designating definite descriptions for corresponding proposition-designating descriptions (Harman [1]). To slightly modify one of Harman’s examples, “The fact that there was an explosion made the basement burn” doesn’t seem to entail “The (true) proposition that there was

[10] Both Hirsch’s work on verbal disputes and his work defending common-sense ontology contain a number of hypotheses about constraints on interpretation that may be used to undermine step (2).

an explosion made the basement burn.” Furthermore, many hold that there are no non-obtaining facts, while there are false propositions.

Second, there is some relatively familiar natural language evidence against individuating facts so coarsely. I’ve already mentioned anaphoric pronouns that appear to designate both facts and propositions. Notice that such pronouns appear to provide the objects of propositional attitudes, so any motivation for individuating such objects hyperintensionally may extend to facts. For example: “Snow is white. That’s a fact, and I have always known it.”

Similarly, descriptions for necessarily co-instantiated facts cannot be freely intersubstituted *salve veritate*. From the truth of “The fact that there was an explosion made the basement burn”, it does not follow that “The fact that there was an explosion and $2 + 2 = 4$ made the basement burn”.

Third, the complicated nature of the linguistic evidence may lead one to claim that “fact” is polysemous, designating multiple different kinds of entities. Such a view doesn’t obviously sit well with a fact-centric motivation for the relevant partitioning principle. On such a view, the notion of a fact is one of the central notions to determination of metaphysical merit. If there are multiple types of facts, we’d like to understand why we’re focusing on a particular one.

I’ve run through these problems relatively quickly. The point is not to reach a clear verdict on the version of the equivalence-class strategy being pursued. Rather, it was to bring to light some deeper issues that underlie this particular development of deflationism. I take it to be a virtue of the equivalence-class strategy that it ties the notion of metaphysical merit to independently tractable issues in metaphysics and philosophy of language.

14.3.2.3 Modality and Deflationism

A deflationist who adopts an intensional view of facts should reject a deflationary view about modality. The reason is that if they don’t, the result will be that disputes about which disputes are deflated will themselves be deflated. To see this, it will help to assume a fairly specific account of possible worlds. Take them to be long conjunctive propositions: intuitively, propositions that, in some sense, describe a maximal scenario. Now consider a dispute about whether a particular conjunctive proposition describes a possible world or not. The deflationist will take there to be multiple candidate interpretations of “possible”, at least one vindicating each disputant, such that no candidate interpretation has more metaphysical merit.

If this sort of deflationism about possibility is adopted, it will follow that questions about intensionality are themselves at least partly deflated. Whether a sentence S is intensionally equivalent to a sentence S’ may depend on whether proposition R is possible. This question, in turn, may be deflated. Since the question of intensional equivalence is deflated, the question of metaphysical merit itself will likewise be deflated.

Some deflationists may react by embracing a deflation of deflationism itself. However, I think this would be ill-advised. One of the thoughts driving the sort of deflationism that we're considering is that ordinary object ontological debates are exceptional in their shallowness. That is to say that there is a genuine contrast between debates over the existence of tables and debates in mathematics. The deflationist about deflationism will take this distinction itself to be shallow.

I don't have an argument for or against the deflationist about ordinary object ontological dispute also adopting a substantive view of modality. As a matter of fact, however, I have encountered numerous philosophers who seem attracted to deflating both sorts of debates. Such philosophers will have to find a non-modal way to defend of their position.

This brings to light another feature of deflationism more generally. The deflationist who pursues the go-to move needs some way to make sense of the notion of parity or metaphysical merit. I've provided them with one such way: to induce an equivalence-class on propositions that tracks merit. Defending such a partitioning itself requires some substantial metaphysics. This becomes especially clear if, as I've suggested, the deflationists tie their partition to a view about the individuation of facts. The fact that deflationism about local matters requires an inflationary position about others is a point repeatedly stressed by Sider [5].

14.3.3 Deflationism and Partitioning Principles

As I've been stressing, none of these worries decisively rule out pursuing the equivalence-class strategy with intensional partitions. However, they have given us reason to doubt both its motivation and tenability. A natural question then arises: can we vindicate deflationism about ordinary object ontological disputes with another version of the equivalence-class strategy? A completely satisfying answer to this question will require evaluating the prospects for a variety of distinct partitioning principles. That is certainly beyond the scope of this discussion. However, it is worth making some remarks.

First, recall that we may attempt to defend a particular partitioning principle in one of two primary ways: by linking it to a general ontology of facts, or by linking it to more specific metaphysical considerations. In introducing this distinction, my example of the second sort was a tempting thesis about the link between converse relations and facts: that the instantiation of a relation by two objects corresponds to the same fact as instantiation of its converse by the same objects (in reverse order). Despite making this distinction, I have yet to seriously consider any partitions motivated by such specific considerations. However, a proponent of QV that wants to deflate ordinary object ontological disputes can reasonably appeal to such considerations. Insofar as we're tempted by the fact-identity claim for converse relations, we may be tempted by a similar claim for constitution. Perhaps facts concerning the arrangement of some constituting matter are identical to facts concerning the arrangement of the constituted matter. If that's the case, then we may have some reason to think that "There are$_{csl}$ tables," and "There are$_{csl}$ simples arranged tablewise," express propositions in the same partition.

Second, and relatedly, note that a would-be deflationist need not partition the entire set of propositions. It would suffice to deflate the debate in question that the opposing sentences uttered in the dispute belong to the same cell in the partition. This limited claim can be combined with agnosticism about the more general partitioning principle.

Third, recall that the friend of QV need not be an enemy of the claim that some hyperintensional distinctions are metaphysically important. This is hinted at when Hirsch disavows rejection of worldly quantificational structure. Insofar as QV is combined with such a view, it would be natural to link such hyperintensional distinctions to partitioning principles. Perhaps, for instance, a proposition is in the same cell as any conjunctive proposition expressing its complete grounds. The important point is that deflationary opportunities may actually be discovered by combining QV with hyperintensional views of facts and dependence.

14.4 The Prospects for Deflationism

The goal of this paper was not to defend deflationism about any particular dispute or undermine it. Rather, I set out to investigate the relationship between deflationism and intensionality in regards to metaphysical merit/parity.

In the realm of linguistic and mental content, I argued that the deflationist should not adopt an intensional view of content. Once this was established, I outlined what I take to be a more promising route to understanding metaphysical merit: the equivalence-class strategy. I then examined whether such a strategy could be utilized to link merit to intensionality. There are a number of key challenges to establishing such a link. However, even if the link can't be established, not all is lost for the deflationist. The equivalence-class strategy can be pursued in numerous different ways, supporting numerous different deflationary conclusions. Furthermore, utilizing such a strategy allows the deflationist to connect merit to numerous independent issues in metaphysics, ensuring that a claim about metaphysical merit need not be obscure or unsupported.

References

1. Harman, G. 2004. Category mistakes in M&E. *Philosophical Perspectives* 17(1): 165–180.
2. Hirsch, E. 1993. *Dividing reality*. Oxford: Oxford University Press.
3. Hirsch, E. 2011. *Quantifier variance and other essays*. Oxford: Oxford University Press.
4. Sider, T. 2009. Ontological realism. In *Metametaphysics*, ed. D. Chalmers, D. Manley, and R. Wasserman. Oxford: Oxford University Press.
5. Sider, T. 2012. *Writing the book of the world*. Oxford: Oxford University Press.
6. Stalnaker, R. 1978. Assertion. In *Syntax and semantics*, vol. 9, ed. P. Cole. New York: Academic.
7. Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT.
8. Urmson, J.L. 1956. *Philosophical analysis*. Oxford: Oxford University Press.

Part IV
Metaphysics and Ontology

Chapter 15
Making Quantified Truths True

Axel Arturo Barceló Aspeitia

Abstract In this paper, I present a novel way of meeting the challenge of grounding the truth of contingent quantified truths in positive, actual aspects of reality. My solution recovers the commonsensical intuition that what makes <all *A*s are *B* > true (in a circumstance of evaluation *w*) is just the *A*s (in *w*) being *B* (in *w*). The proposal is based on recognizing that the metaphysical relation that binds truths to their truth-makers is defeasible. Consequently, it is possible for a truth-maker to make a truth-bearer true in some circumstances of evaluation, but fail to do so in others, in those others where appropriate defeaters exist.

Every truth tells us something about the world, that is, about how the world *is*. This much is not controversial, as it is not controversial to claim that every truth is true because of something in reality that—so the expression goes—*makes* it true; in philosophical lingo, this means that in every possible world, situation, or circumstance of evaluation where a truth-bearer is true, that truth is grounded in some determinate truth-maker existing in that possible world situation or circumstance of evaluation.[1] Yet there is no one-to-one correspondence between truth-bearers and truth-makers, since some truth-bearers can be made true by different truth-makers in different worlds and some truth-bearers are not made true by any truth-maker in any world. As a matter of fact, it is commonly assumed that for basic synthetic singular truth-bearers that predicate a property or relation of one or many objects—truth-bearers like <Curiosity is in Mars> or <Peter drives a Mazda>—whatever truth-maker makes a truth true in one world also makes it true in any other world (where both truth-maker and truth-bearer exist and, therefore, where the respective

[1] In what follows, for simplicity and without loss of generality, I will speak only of possible worlds.

A.A. Barceló Aspeitia (✉)
Instituto de Investigaciones Filosóficas, UNAM, Circuito Mario de la Cueva, Ciudad Universitaria, Del. Coyoacán, México D.F. 04510, Mexico
e-mail: abarcelo@sofia.filosoficas.unam.mx

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_15

truth-bearer is true).[2] Yet, this is not true for other sorts of truth-bearers, like disjunctions or existentials. For example, different truth-makers, each involving different people, may make the truth-bearer <Some people like yogurt> true in different worlds. It is also widely accepted that the same truth-maker might make different truth-bearers true, even in the same world. For example, in a world where the only person who came late to class tonight was John, <John came late to class tonight> and <Someone came late to class tonight> would share the same truth-maker. However, the converse is commonly not accepted, that is, that the same truth-maker might make different truth-bearers true in different worlds, or may make one truth-bearer true in some worlds but fail to make it true in others. On the contrary, one of the most widely accepted principles of truth-maker theory is necessitation, that is, the principle that for every truth-maker f, truth-bearer p, and possible world w, if f makes p true in w, then p must be true in every world w where f exists [11, 12]. Unfortunately, embracing necessitation (in combination with two other fundamental and widely, but not universally accepted principles of truth-maker theory: *truth-maker maximalism*—the thesis that all truths are made true by some truth-maker—and *truth-maker actuality*—the thesis that all truth-makers are positive and actual, that is, entities that exist in reality, rather than absences, lacks, totalities, possibilia, or similar posits) presents its own difficulties for truth-maker theory. Accepting necessitation (plus maximalism and actuality) makes it necessary but very difficult to find adequate truth-makers for a wide variety of truths, like negative, general, and modal truths ([8, 33], etc.).[3]

I take it that this basic problem of truth-maker theory originates from a fundamental asymmetry between truth-bearers and truth-makers. While truth-bearers can be composed of other truth-bearers in several ways—by disjunction, negation, conjunction, quantification, etc.—truth-makers can only be composed in a straightforward aggregative way. In other words, while truth-bearers can be disjunctive, negative, quantified, etc., truth-maker theory resists the existence of negative, disjunctive, or quantified truth-makers. Thus, truth-maker theory must find adequate non-disjunctive truth-makers for disjunctive truth-bearers, positive truth-makers for negative truth-bearers, non-quantified truth-makers for quantified truth-bearers, non-modal truth-makers for modal truth-bearers, etc.

Each one of these problematic kinds of truths presents its own challenges to truth-maker theory. In keeping with the subject of this special volume, in this chapter I will focus exclusively on meeting the challenges that quantified truths raise for truth-maker theory. In particular, I will propose and provide a motivation for a conception of truth-making as defeasible, showing how it deals with quantified truths in general

[2] In this paper, I adopt the widespread convention of surrounding truth-makers in square brackets and truth-bearers in less-than and more-than symbols.

[3] In the context of this paper, therefore, I will call truth-maker theory the theory that aims at specifying in a systematic way what portions of the world make what truths true. Other theorists (for example, Fine [15, 16]) have used the same term to refer to other theories pursued with other purposes, and my proposal may fail to achieve some of those.

and contingent general truths in particular. Certainly, I think this proposal gives us a general way of dealing with similar problems for other kinds of truths, but I shall leave the defense of that stronger thesis for a later occasion.

It is a fairly well known fact in the truth-maker literature that if one rejects necessitation, finding adequate positive truth-makers for general truths is fairly straightforward and easy [34]. However, rejecting necessitation just because it is convenient cannot be enough reason. We need a principled account of truth-making to ground a rejection of necessitation. To this end, I will argue that people who defend necessitation are mistaken about the role of truth-makers in metaphysically explaining truth. I will defend the view that the fundamental theoretical role of truth-makers is to explain why true truth-bearers are actually true, independently of whether or how they could have been true or false in other, counterfactual situations. From this perspective, one can see why the truth-making relation ought to be defeasible, for whatever explains why a given truth-bearer p is actually true (i.e., in the actual world) might not explain why the same truth-bearer p could be true in a different world, even if its truth-maker also exists there. Consequently, once we recognize truth-making as a defeasible relation, we can do away with the principle of necessitation and provide adequate truth-makers for all actual truths.

The paper will proceed as follows. In the first section, I will present the well-known problem of providing truth-makers for contingent general truths. Then, I will introduce my proposal of truth-making as a defeasible relation and show how it deals with the aforementioned problem. Finally, I will present a couple of arguments in favor of necessitation already in the literature and show how they all fail against my conception of truth-making as defeasible.

15.1 Quantified Truth

If the existential truth-bearer <There is an A > is true in a world w, then there is an object a that is A in w. Since this a exists and it is A in w, then the positive truth-maker [a is A] makes <There is an A > true in w. This much is not controversial. Nevertheless, quantified truths of other sorts present serious problems for truth-maker theory. Here is a simple example: Consider the truth-bearer <My nephews are mischievous>.[4] Intuitively, it would be made true by the truth-maker that [Xaime and Balam are both my nephews and are mischievous] (or by the aggregate of the truth-makers [Xaime is my nephew], [Balam is my nephew],[5] [Xaime is

[4]The problem is common to all general truth-bearers that are accidentally true. Cf. Rosen [40]. For how to deal with non-accidental generalization, see Fine [15, 16], Lewis [25], Armstrong [1].

[5]Since being a nephew is a relational property, it is very likely that what makes true that Xaime is my nephew does not only involve her, but also me, her mother (my sister), and the kinship relations between us. In order to simplify my exposition, I will ignore this complication for the rest of the article.

mischievous] and [Balam is mischievous]). However, this truth-maker (or aggregate of truth-makers) does not satisfy necessitation. Consider a possible world, as similar to this world as possible, but in which my brother had a previous child called Piero. In that possible world, Piero is very, very well behaved and not mischievous at all. In such a world, Xaime and Balam are still my nephews and still as mischievous as in this world (thus, the relevant truth-makers that putatively make the truth-bearer true in this world, would also exist in such a world); however, in that world, it is no longer true that my nephews are mischievous. Thus, we have to conclude, that such truth-maker cannot make the relevant quantified truth-bearer true. Something else, another truth-maker, seems to be missing, something that makes it true that I have no other nephews besides Xaime and Balam. But what truth-maker could this be? It seems that it cannot be just that Xaime and Balam are my nephews, for once again this truth-maker does not necessitate that there are no other nephews besides them. So it seems like a truth-maker of a very different sort is required: a negative truth-maker (that I have no other nephews), a totality truth-maker (that these are all of my nephews), or something else.

Thus, we are faced with four options: we either (i) embrace truth-makers of a different sort (totality or negative truth-makers) ([3, 4], etc.), (ii) find that elusive positive truth-maker that, together with the truth-maker that both Xaime and Balam are my nephews and are mischievous, does necessitate the truth of the truth-bearer that <My nephews are mischievous> [13], (iii) reject the principle of necessitation [34], or finally, (iv) accept that some truths have no truth-makers [5, 29–31].

In this paper, I want to defend option (iii), but I will adopt a different strategy than usual. Usually (for example, [6, 7, 28, 34, 35]), advocates of (iii) do not defend it directly, but instead argue against the other three options outlined above, that is, against accepting odd truth-makers, truths without truth-makers or there being some undiscovered positive truth that does necessitate truth-bearers of this kind. In this paper, in contrast, I will try to defend (iii) directly. First, I will outline an account of truth-making that motivates rejecting necessitation and then use such account to defuse those arguments in its favor that have recently surfaced in the literature.

15.2 Defeasible Truth-Making

Before determining whether or not truth-making is a defeasible relation or not, it is important to be clear first on what exactly makes a relation defeasible. Before entering the philosophical vocabulary, "defeasibility" was a technical legal term referring to those estates or interests in land that can be rendered void by certain given circumstances [17]. It was later extended by Hart [19] to cover all legal concepts and relations that can also be rendered void by external circumstances. As Hage further explained:

> The concept of a contract is a typical example. A contract that has come into existence after an offer and acceptance can be invalidated if one of the parties involved invokes a defeating condition, such as fraudulent misrepresentation, or undue influence. In this connection it is

> crucial that the defeating conditions are actually invoked; the mere fact that they occurred is not sufficient to defeat the contract. Therefore, defeaters are to be distinguished from ordinary conditions for the existence of a contract, which do not need explicit invocation. [17, p. 222]

Today, the notion of defeasibility has become commonplace in the philosophical landscape, bearing fruitful applications in logic, ethics, epistemology, metaphysics, etc. In logic, for example, it has proved very helpful in modeling a special kind of good reasoning, where the conclusion is nevertheless not necessitated by the premises, but can be *defeated* by extra information. Consider the following example, adapted from Prakken and Vreeswijk [37]: A newspaper editor might decide not to publish information about the health of a public figure on the grounds that such information is not of public interest and to publish it would be a violation of the person's right to privacy. However, if she finds out later that the public figure is about to run for public office, she might change her mind on making such information public. If she thinks that the politician's health might affect his performance as public officer and that this information might help public deliberation regarding matters of public concern, the editor might now withdraw her previous conclusion and publish the information. Notice that, before she found out about the subject's plans of running for office, the editor's reasoning was *not* flawed at all. She had good reasons for reaching the conclusion she reached. Yet, the information she possessed did not necessitate the truth of such conclusion; that is why she could rationally retract from her conclusion after acquiring new information. Her original reasoning was good, even if defeasible. It was defeasible because it could have been, and actually was, defeated by the acquisition of new information. Appealing to the defeasibility of her original reasoning allows us to explain why it was rational for her to change her mind and not stick to her previous conclusion, while also recognizing what was right about it. In general, this is how defeasibility, as a notion, can help us to better understand relations that are not necessary, yet are still valid in the absence of defeating circumstances.

Consider another example, this time from Chisholm [9, 10]. According to Chisholm and other epistemologists (Pollock, Plantinga, etc.), the evidence sensory appearances give us for believing in basic facts about the physical world are defeasible. For example, if I have the sensory experience as of being in the presence of something red, then I am right in believing that I am actually in the presence of something red, for such sensory experience gives me pertinent good evidence. However, this evidentiary relation can be defeated, if, for example, I learn that my environment is abnormal in a relevant sense (for instance, all the ambient light is red) [24]. Notice that for my experience to give me good enough reasons for my beliefs, I need not have the extra belief that my environment is relevantly normal (and that the ambient light is not red). The absence of defeating circumstances is not *another* reason for having the appropriate belief, it is something external that does not figure among the relata of the relevant epistemic relation. The evidential relation holds only between the experience and the belief, not between the experience *plus the absence of defeating circumstances* and the belief. When there are no exceptional defeating circumstances, the epistemic relation holds *fully* between experience and belief; it does not hold in some diminished or incomplete sense. In the absence of defeating circumstances, defeasible evidence is good enough evidence; it is as good evidence as any, even if it does not necessitate the truth of the belief it supports.

In general, to say that a relation is defeasible is not to say that it does not hold, or that there is something missing from it. In the absence of the relevant defeaters, the defeasible relation holds completely and in full force. Consider Hart's original example of a legal contract. In the absence of any defeating circumstances, a legal contract is fully binding. What binds those who sign a contract is just the contract, not the contract *plus the additional condition* that no defeater takes place. The absence of defeating circumstances is not part of what binds the signees. The existence and signing of the contract on the one hand, and the absence of defeating circumstances on the other, are not on a par in this regards. As a matter of fact, one of the main theoretical purposes Hart had for introducing the notion of defeasibility into the philosophy of law was precisely to introduce a fundamental distinction between two substantially different kinds of facts involved in legal binding phenomena: those positive facts and actions that are fully responsible for making the relevant legal relation binding, and those other facts and actions that could nullify it. It is my claim that a similar distinction needs to be drawn in metaphysics between those conditions that are fully responsible for making a truth-bearer true and those conditions that can make the relation between truth-maker and truth void. Just as it is a mistake to consider the absence of defeating circumstances like skeptical scenarios as part of the reasons behind our empirical beliefs, it would be a mistake to consider the absence of certain facts as part of what makes any truth true.

In every case where the notion of defeasibility has been helpful, it has been as a theoretical tool for dealing with relations where, if aRb (i), the existence of a (and b) does not entail that aRb, and for every pair a and b such that aRb, there is a class of (possibly mutually inconsistent) defeaters $D = \{d_1, d_2 \dots d_n\}$ (for some $n > 0$) such that (ii), if a and b exist, $\sim aRb$ if and only if any of them is the case, so that (iii) a and the absence of $d_1, d_2 \dots d_n$ are not R-related to b. It is my claim that truth-making is a relation of just this kind: if a truth-maker f makes a truth p true, (i) the existence of the truth-maker f does not necessitate the truth of p, and for every pair of relata f and p such that f makes p true, (ii) there is a class of (possibly mutually inconsistent) defeaters $D = \{d_1, d_2 \dots d_n\}$, such that, even if f exists, f does not make p true if any defeater d in D is the case, so that (iii) it is not f and the absence of the defeaters $d_1, d_2 \dots d_n$ that makes p true. In other words, truth-making is a defeasible relation. This means that every time a truth-maker f makes a truth-bearer p true, there is a class of (possibly mutually inconsistent) defeaters $D = \{d_1, d_2 \dots d_n\}$ for the truth-making relation between p and f, such that among the worlds where f exists, f makes p true in every world and only in those worlds where none of the facts in D exist.

Now, before defending my account of truth-making as defeasible, I want to clean up the conceptual space by stressing a basic feature of truth-making whose importance I think has been under-appreciated: It should be noted that truth-making is primarily a relation between truths and what makes them true, that is, between truths and their truth-makers, not between truth-bearers and truth-makers. This must be obvious from the fact that the same truth-bearer may be made true by different truth-makers in different possible worlds. In contrast, every truth has its own specific truth-maker, since for every truth-bearer there is a single specific truth-maker that makes it true in every possible world in which it is true. Even though truth-maker theory must give a systematic account of why and which different truth-makers make the same truth-bearer true in different possible worlds, each truth-maker is responsible for grounding the truth of its truth-bearer only in its particular and proper world, and not in others. This means that whatever makes a truth-bearer p true in a possible world w cannot be considered at fault for not responding to whatever happens to p in a different world w', especially if p is not true in w'. In other words, the goal of truth-maker theory is to tell us why what is true is true, but not why what is not true is not true.

Under the conception of truth-making as defeasible that I advocate, truth-makers are positive aspects of the world that ground truth, but they need not necessitate the truth they ground because there may be other events taking place in other possible worlds that prevents the truth grounded in this world from extending to those other worlds. Thus, just as in law, epistemology and logic, we can divide whatever strictly necessary conditions there are for the truth of a truth-bearer between those positive aspects that actually make a given truth true and those external circumstances that might defeat that truth-making relation. The former, positive aspects of reality are the possible truth-makers, while the latter, external circumstances, are its possible defeaters. The truth-making relation holds only in the absence of the defeaters. If they are present, the truth-making relation is defeated and the relevant truth bearer is no longer true and, hence, there is no truth to make. In the absence of defeaters, however, the first ones are fully responsible for making the relevant truth true, and no mention of the absence of the seconds is necessary. Notice the strong analogy between the truth-making case and Chisholm's epistemic one described above. In both cases, the absence of defeaters should not be considered among the relata of the relevant relation. In our case, this means that the absence of defeaters is not part of what makes any truth true. Even if the existence of a truth-maker f is sufficient for the truth of a given truth-bearer p only in those cases where no defeater d exists, we still have good reason to say that it is f that makes p true instead of saying that it is f plus the absence of the defeaters d that together make p true.[6] The absence of defeaters does not actually play any role in making true truth-bearers true and thus should not be considered part of any truth-making aspect of reality. Drawing the distinction between truth-makers and defeaters allows us to do justice to the intuition that a true truth-bearer p's truth-maker is whatever exists in the actual world that makes p true.

[6]Contrast this with the correct principle that if a were strictly sufficient to necessitate the truth of p only when b exists, where a and b are both positive entities, then it would actually be the aggregate of a and b what makes p true.

Adopting this view of truth-making as a defeasible relation can help us solve the above problem of my mischievous nephews. If you recall, the challenge was to explain why [Xaime is my nephew], [Balam is my nephew], [Xaime is mischievous] and [Balam is mischievous] did not seem enough to make <My nephews are mischievous> true in the actual world, where Xaime and Balam are my only nephews. This was supposed to be a challenge because, even if I actually do not have any further nephews, the mere possibility of me having some well-behaved nephew makes it possible for the aforementioned facts to exist in counterfactual circumstances where the relevant generalization is false. However, if we take the truth-making relation to be a defeasible one, we can see that no extra truth-maker in the world is necessary to make the quantified truth-bearer true. The defeasible view allows us to draw a sharp distinction between actual facts about my actually mischievous nephews and merely possible facts about possibly well behaved nephews. The first ones are the aspects of the real world that are fully metaphysically responsible for it being true about the *actual* world that my nephews are mischievous. They are the ones that make the quantified truth-bearer true. The other merely possible facts are in no way involved in making this truth true. Instead, they are the possible defeaters of the truth-making relation between those facts and the actual truth of <My nephews are mischievous>. Just as it is not the sensual experience plus the absence of defeating circumstance that epistemically supports an empirical belief, it is not that Xaime and Balam are my nephews and are mischievous plus the absence of defeating circumstance that metaphysically grounds the truth of <My nephews are mischievous>. The relevant facts about my actual mischievous nephews are good enough to fully make the quantified truth-bearer true. Once Xaime and Balam are my nephews and are mischievous, there is nothing else missing in order for my nephews to be mischievous.

So, what happens in that other world where Piero is so well-behaved? There, the relevant truth-makers also exist, but they do not make the truth-bearer true because there is an object in that world that defeats the truth-making relation: my nephew (in that world) who is not mischievous (in that world). But all of that does not happen in the actual world and thus, is irrelevant to what is *actually* true, that is, to what is true in the actual world. After all, whatever is the case in other worlds bears no relevance whatsoever to what makes truth-bearers true in the actual world, and it is only this that truth-maker theory must account for. In other words, the existence of Piero as a well-behaved nephew of mine works as a defeater of the truth-making relation between the actual truth of <my nephews are mischievous> and its (actual) truth-makers, and that explains why, even though my *actual* nephews are still mischievous *in that possible world*, it is not true that my nephews *in that possible world* are mischievous.

In general, my proposal is thus: what makes any contingently general truth of the form "All the *X*s are *Y*" (actually) true are just the (actual) *X*s being *Y*. No mention of any other merely possible *X*s is necessary to account for this truth. Any other possible *X*s that are not *Y* are just possible defeaters of this truth-making relation; in their absence, the *X*s that are *Y* are grounds enough for the truth of the generalization. In the case of the truth-bearer <My nephews are mischievous>, for every object x in the actual world that is one of my nephews in the actual world the (positive and actual) truth-maker that x is mischievous exists in the actual world. Since all the objects in the actual world that are my nephews in the actual world are Xaime and Balam, this means that all that is necessary for the truth of the truth-bearer <my nephews are mischievous> is that Xaime and Balam be mischievous and nephews of mine.

Note that a corollary of adopting a view of truth-making as defeasible is that we have a principled way of rejecting the validity of the principle of necessitation. That a truth-maker f makes a truth-bearer p true in the actual world is completely compatible with the existence of a different world w where f also exists but p is not true, as long as at least one defeater d for the relation between f and p exists in that world w. This way, the introduction of a truth-making defeater is not simply and ad-hoc rejection of necessitation. However, even if adopting a view of truth-making as defeasible gives us a principled way of rejecting the principle of necessitation, it is still necessary to reply to those arguments in the literature in favor of this principle. That is the goal of the reminder of the paper.

15.3 Against Necessitation

> But a truth-maker that does not necessitate the truth of the truth-bearer it makes true, sounds awfully like a truth-maker that does not make its truth true, that is, a truth-maker which is not a truth-maker. [8, p. 264]

Many people before me have challenged necessitation and, while in the early years of truth-maker theory necessitation was mostly an unchallenged assumption, some arguments in favor of necessitation have surfaced recently. Unfortunately, in the few places in the recent literature where necessitation has been discussed, many philosophers have found that necessitation is obvious and does not need to be justified (for example, [8]), or have tried to derive it from what they consider more basic principles of truth-making, like:

Truth-Maker Essentialism: Every truth has a truth-maker, *which is essentially that truth's truth-maker*. (Parsons [34], emphasis in the original), or

Truth-Maker Sufficiency: The existence of a truth-bearer's truth-maker must be sufficient in itself for such truth-bearer to be true [2]

The logical relations between these three principles is a contested matter [21], but it is commonly assumed that truth-maker necessitation follows more or less directly from either of them. However, they are hardly any more intuitive or widely accepted than necessitation itself, so appealing to them would not convince anyone not already convinced of necessitation. Thus, we can easily dismiss these arguments and move on to a couple of other more substantial arguments that also try to deduce necessitation from other more fundamental principles of truth-making:

Truth-conditional Essentialism: It is essential to truth bearers to have the truth-conditions they have [18], or

Truth-Making as Explanation: The existence of the truth-maker must explain the truth-bearer's truth.

In what follows, I will try to explain why somebody might think that accepting any of these principles commits us to embracing necessitation, and I will show why, in adopting a account of truth-making as defeasible, it is possible to accept any of them without having to accept necessitation.

15.3.1 Truth Conditions Are Constitutive of Truth-Bearers

It is commonly assumed that truth conditions are metaphysically necessary, that is, that it is necessary for any truth-bearer to have the truth conditions it has (this is implicit in the sets-of-possible-worlds conception of truth-bearers, for example), and therefore that it is necessary for any truth-bearer to be made true by the truth-makers that make it true [18, 41]. This latter thesis, of course, is nothing but necessitation; thus one way of challenging necessitation is by challenging the widespread assumption that every truth-bearer has some fixed truth conditions that do not vary from circumstance of evaluation to circumstance of evaluation. Andrea Iacona makes the point this way:

> Just as for a [truth-bearer] to be true in the present and actual state of affairs is for its truth condition to be satisfied in that state of affairs, for a [truth bearer] to be true in a state of affairs different from the present and actual one – call it an "alternative" state of affairs – is for its truth condition to be satisfied in that state of affairs. [20, p. 38]

And here is the same point made by Hanks:

> But it can't be an accident that this [truth-bearer] has these representational features; it wouldn't be that very [truth-bearer] if it somehow had different truth-conditions. This means that a [truth-bearer] cannot have its representational features conferred on it by something external, on pain of making those features inessential to the [truth-bearer]. Anything worth calling a [truth-bearer] must have its truth-conditions on its own, in the sense that its truth-conditions arise out of its internal make-up. [18, p. 474].

Recently, Richard [38], MacFarlane [26, 27] and others have challenged this assumption. They have defended views where there are genuine truth-bearers whose truth conditions are not fixed, but vary from context to context, so that for example, the same truth-makers about carrot cake can make true the truth-bearer <carrot cake is delicious> in some contexts, but not in others, while the truth-maker [the 15th of June, 2012 is Friday] can make the truth-bearer <today is Friday> true in certain contexts (contexts where today is the 15th of June) but not in others. Soames [44] and King [22, 23] have also rejected the thesis that truth-bearers have the truth conditions they have by necessity and instead maintain that they have them in virtue of contingent matters of use or intentionality. Thus, if any of Richards, MacFarlane, King, or Soames is right (and this is a big 'if'), truth conditions would not be essential to truth-bearers.

As sympathetic as I am to some variation of relativism, this is not the place to mount a full on attack on the claim that it is essential for truth-bearers to have the truth conditions they have; instead I will argue that truth-conditional essentialism does not actually entail necessitation, since the same truth-conditions can be met by different truth-makers in different circumstances of evaluation. This is partly what a defeasible understanding of truth-making amounts to, after all. Under an understanding of truth-making as defeasible, the necessary (and sufficient) conditions for making <my nephews are mischievous> true are essentially the same in any circumstance of evaluation, to know: that my nephews be mischievous. This is the same in any possible world, independently of who my nephews might be. In general, any truth-bearer of the form <all fs are g > is true in a possible world w if and only if all the fs in w are g in w. In any possible world, for the truth-bearer to be true, this same condition must be met. That is why a conception of truth-making as defeasible does not violate truth-conditional essentialism. Truth conditions remain constant from possible world to possible world. What can change, and sometimes does, is which truth-makers are involved in satisfying those same conditions in different circumstances of evaluation. In our actual circumstances, for example, the truth-makers involved in satisfying the condition that my nephews be mischievous are [Xaime is my nephew], [Balam is my nephew], [Xaime is mischievous] and [Balam is mischievous]. In a different circumstance of evaluation, for example, one in which I had another mischievous nephew, Carlos, besides Xaime and Balam, this *same* truth condition would be satisfied by the existence of other truth-makers: [Xaime is my nephew], [Balam is my nephew], [Carlos is my nephew], [Xaime is mischievous], [Balam is mischievous] and [Carlos is mischievous]. What it takes for the truth-bearer <my nephews are mischievous> to be true in this circumstance of evaluation would be the same as what it takes for the same truth-bearer to be true in any other circumstance of evaluation, that is, that all of my nephews be mischievous. The truth-makers that satisfy these conditions might indeed change, but that is just to be expected in cases like these, given the defeasible nature of the truth-making relation.

15.3.2 Truth-Making Must Explanation Truth

Pagès [32] has argued that necessitation is necessary for any explanatory relation; in other words, that the explanandum of any explanation must be the case in any possible world or circumstance of evaluation where the explanans is the case, and that since truth-makers are supposed to explain why true truth-bearers are true ([40, 42], etc.), they must also necessitate the truths they explain.

In response to Pagès' argument, Pendlebury has challenged the claim that explanation requires full-fledged necessitation:

> An explanans must, at some level, necessitate its explanandum; it is not enough that both just happen to hold. But it is, I think, enough if the explanandum holds in all cases in which the explanans holds that are sufficiently similar to the actual case. This would secure the

> important result that the explanans establishes that the explanandum was to be expected, which I take to be necessary for an adequate explanation. [36, p. 142][7]

The main point behind Pendlebury's counter-argument is that explanations depend on expectations, and what is expected or unexpected is a very context-sensitive matter. Thus, explanations are also very context sensitive; that is, something might be a good explanation for something in some circumstances but not in others [46]. However, once we recognize the context-sensitivity of explanations, we can easily see that explanation does not require necessitation. For Pendlebury, for *a* to explain *b*, it is not necessary that *a* holds in all the worlds or circumstances of evaluation where *b* exists, but only in the closest ones. This means that, for example, for the truth-maker that Xaime and Balam are my nephews and are mischievous to explain the truth of the truth-bearer <My nephews are mischievous>, it is not necessary that this truth-bearer be true in every circumstance of evaluation where the truth-maker exists, but only in those relevantly close to the actual one, and since presumably those circumstances where I have a third or fourth nephew are more dissimilar to the actual one than those where Xaime or Balam are not mischievous, the relevant conditional holds: if Xaime and Balam are my nephews and are mischievous then my nephews are mischievous.

I think Pendlebury is right in saying that explanation does not entail necessitation, because it is too context sensitive. Nevertheless, I do not think his way of developing the argument in terms of closeness among possible worlds or circumstances of evaluation is the best one. Behind Pendlebury's response to Pagès lies one of the most pressing questions in truth-maker theory: what sort of modal relation links truth and truth-maker? It is well understood that simple material implication is too weak a relation, and that strict implication is too strong [39], so Pendlebury suggests using a *subjunctive* conditional. According to Pendlebury, if t makes p true in a given world w, then the subjunctive conditional "if t exists, then p exists too" must hold in w as well, where a conditional of this sort is true in a possible world w if p is also true in all the worlds closest to w where t also exists. The problem with Pendlebury's account is that if t makes p true in a given world w, then the subjunctive conditional "if t exists, then p exists too" is *trivially* true in w, and thus, this conditional is also too weak to do the job. If t makes p true in w, then t exists in w, and thus the closest circumstance of evaluation where the truth-maker t exists is just w itself, and in that world, by hypothesis, the truth-bearer p is true. Also, notice that requiring that the subjunctive conditional be true in every possible world, instead of only in w, would not work either, since that would be equivalent to necessitation, which is just what Pendlebury wants to avoid.

Thus, a better (but not necessarily the only) way of presenting the same idea is to avoid talk of closeness and instead notice that explanations always involves some sort of contrast. When we ask for an explanation of something, we are always asking why that happened *instead of something else* that was expected or otherwise

[7] Schnieder [42] makes essentially the same point.

contextually salient. Let's take an example. Suppose that John is sitting in the hall. Without a proper context, it is hard to understand the question "Why is John in the hall?" One might be expecting an explanation as to why *John*, instead of Robert or Peter, is in the hall, or why is he *in the hall*, instead of the lobby, or why is he *still* in the hall and not yet gone, etc. Not all the information relevant for one of these explanations might be relevant to the others. The explanation of why John is in the hall instead of the hallway might be that our secretary let him in, but that might not explain why he is still there or why he is not at the bank or the post office. In this way, the process of explaining always involves a selection of information relevant to the contrast in question.

The problem for those who expect explanation to entail necessitation is that once we start selecting information, we are also *losing* information, and once we lose some information, we also lose necessitation. That the secretary let John into the hall might explain why John is in the hall (rather than the hallway) but does not necessitate it, because, for example, he could have left already. Thus, explanation does not entail necessitation.

In order to avoid this problem, one might be tempted to suggest that truth-making must not only explain why something is true, but must *completely* explain it [14], so that for a truth-maker to ground a truth, it must explain it in *every possible context.* This suggestion, however, is far from satisfactory. For starters, it is blatantly ad-hoc. There does not seem to be a principled reason to demand complete explanation instead of simple, normal explanation. Furthermore, requiring a complete explanation is too demanding and would go against our most basic intuitions of how truth-making works. If we demanded complete explanation for truth-making, then it would no longer be true that, for example [John is sitting in the hall] would make <John is sitting in the hall> true, since to fully explain why John is sitting in the hall, in some circumstances it might be relevant to explain why it is John instead of Robert that is sitting in the hall; but John sitting in the hall does not explain this (after all, that John is sitting in the hall is metaphysically compatible with Robert also sitting there). However, this is a paradigmatic example of truth-making, so whatever sort of explanation is involved in truth-making, it must be between the simple truth-maker [John is sitting in the hall] and the truth of <John is sitting in the hall> [43, 45]. Consequently, complete explanation is too strong a demand on the relation of truth-making.

One might also try to appeal to some *sui generis* notion of metaphysical explanation that does not behave quite like ordinary explanation, and in particular, is not context sensitive and actually requires necessitation between explanans and explanandum. However, such a move would not be advisable to the advocate of necessitation, since it would seriously cripple the case against defeasibility. Notice that the plausibility of the principle of explanation depends on "explanation" being understood in a non-technical way. The more we depart from the traditional understanding of explanation, the less plausible such principle becomes. Demanding a stronger and *sui generis* relation of explanation thus takes this line of reasoning too close to an ad hoc and question-begging position.

15.4 Conclusions

In this paper, I have offered a novel way of solving the old problem of grounding the truth of contingent quantified truths in positive, actual aspects of reality. My solution recovers the commonsensical intuition that what makes a truth bearer of the form <all As are B > true are just facts about the As, that is, that they are B. This solution is based on recognizing that truth-making, that is, the metaphysical relation that binds truths to their truth-makers is defeasible and can be voided in certain circumstances. This means that, contrary what most truth-maker theorists believe, the existence of a truth's truth-maker does not necessitate that truth; that is, it is not the case that, if f makes p true, then p is true in any circumstance of evaluation in which f exists. On the contrary, it is possible for f to make p true, and yet for there to be circumstances of evaluation where f exists, but p is untrue.

I have also gone through a couple of arguments in the literature that have been presented in defense of the principle of necessitation, and I have shown that they mistake the modal aspect of truth-making. First of all, truth-making is a relation that holds primarily between truth-makers and truths, not between truth-makers and truth-bearers. Truths, of course, only exist in worlds; thus, the truth-making relationship is substantially relative to worlds. The notion of truth-makers as defeasible that I have sketched here aims at capturing just those aspects of the (actual) world or circumstance of evaluation that are metaphysically responsible for the truth of the truth-bearers that are *actually* true. So understood, this notion does not entail that whatever makes a truth-bearer true in the actual world must make the same truth-bearer true in any other world, not even in those in which the truth-maker exists. This is so because there exist what I have called truth-making defeaters: facts whose existence has the capacity to defeat the truth-making relation between truth-maker and truth. The existence of such defeaters might explain why some truth-bearers are not true in circumstances where the truth-bearers in question are false, even though their (actual) truth-makers exist, but play no role in explaining why the truth-bearers in question are actually true (i.e., are true in the actual world). Thus, given the well-known problems that result from holding on to necessitation, I conclude that it is better to let go of it and recognize that truth-making is defeasible.

Acknowledgements This project was supported by PAPIIT grants IA400414 *Anti-realismo modal* and IA400412-2: *Epistemología y metafísica de la modalidad.*

References

1. Armstrong D. 1983. *What is a law of nature?* Cambridge: Cambridge University Press.
2. Armstrong D. 1997. *A world of states of affairs*. Cambridge: Cambridge University Press.
3. Armstrong D. 2004. *Truth and truthmakers*. Cambridge: Cambridge University Press.
4. Beall, J.C. 2000. On truthmakers for negative truths. *Australasian Journal of Philosophy* 78: 264–268.

5. Bigelow, J. 1988. *The reality of numbers: A physicalist's philosophy of mathematics*. Oxford: Clarendon Press.
6. Briggs, R. 2012. Truthmaking without necessitation. *Synthese* 189: 11–28.
7. Cameron, R.P. 2005. Truth-Maker necessitarianism and maximalism. *Logique et Analyse* 48: 43–46.
8. Cheyne C., and C. Pigden. 2006. Negative truths from positive facts. *Australasian Journal of Philosophy* 84(2): 249–265.
9. Chisholm, R. 1957. *Perceiving: A philosophical study*. Ithaca: Cornell University Press.
10. Chisholm, R. 1966. *Theory of knowledge*. Englewood Cliffs: Prentice-Hall.
11. David, M. 2009. The correspondence theory of truth. In *The stanford encyclopedia of philosophy*, ed. Edward N. Zalta. (Fall 2009 ed), http://plato.stanford.edu/archives/fall2009/entries/truth-correspondence. Stanford: Stanford University.
12. David, M. 2009. Truth-making and correspondence. In *Truth and Truth-Making*, ed. E.J. Lowe and A. Rami, 137–157. Stocksfield: Acumen Press/Montreal: McGill-Queen's University Press.
13. Demos, R. 1917. A discussion of certain types of negative proposition. *Mind* 26: 188–196.
14. deRosset, L. "Grounding Explanations", Philosophers' Imprint 13: 1–26.
15. Fine, K. 1994. Essence and modality. *Philosophical Perspectives* 8: 1–16.
16. Fine, K. 1995. The logic of essence. *Journal of Philosophical Logic* 24: 241–273.
17. Hage, J. 2004. Law and defeasibility. *Artificial Intelligence and Law* 11: 221–243.
18. Hanks, P. 2009. Recent work on propositions. *Philosophy Compass* 4(3): 469–486.
19. Hart, H.L.A. 1949. The ascription of responsibility and rights. *Proceedings of the Aristotelian Society* 49: 171–194.
20. Iacona, A. 2002. The expressing relation. *Dialectica* 56: 235–260.
21. Keller, P. 2007. A world of truthmakers. In *Metaphysics and truth-makers*, ed. J.-M. Monnoyer, 105–156. Frankfurt: Ontos.
22. King, J.C. 2007. *The nature and structure of content*. Oxford/New York: Oxford University Press.
23. King, J.C. 2009. Questions of unity. *The Proceedings of the Aristotelian Society* 109: 257–277.
24. Koons, R. 2013. Defeasible reasoning. In *The stanford encyclopedia of philosophy*, ed. Edward N. Zalta (Spring 2013 ed.) http://plato.stanford.edu/archives/spr2013/entries/reasoning-defeasible. Stanford: Stanford University.
25. Lewis, D. 1973. Causation. *Journal of Philosophy* 70: 556–567.
26. MacFarlane, J. 2003. Future contingents and relative truth. *The Philosophical Quarterly* 53: 321–336.
27. MacFarlane, J. 2008. Truth in the garden of forking paths. In *Relative truth*, ed. M. Garcia-Carpintero and M. Kölbel, 81–102. Oxford: Oxford University Press.
28. Mellor, D.H. 2003. Real metaphysics: Replies. In *Real metaphysics: Essays in honour of D. H. Mellor*, ed. H. Lillehammer and G. Rodriguez-Pereyra. New York: Routledge.
29. Merricks, T. 2007. *Truth and ontology*. Oxford: Clarendon Press.
30. Merricks, T. 2008. Replies to Cameron, Schaffer, and Soames. *Philosophical Books* 49: 328–343.
31. Mulligan, K., P. Simons, and B. Smith. 1984. Truthmakers. *Philosophy and Phenomenological Research* 44: 287–321.
32. Pagès, J. 2005. Truthmaking and supervenience. *Theoria* 20: 191–197.
33. Pagès, J. 2009. Truthmakers for negatives. *Theoria* 24: 49–61.
34. Parsons, J. 1999. There is no "Truth-Maker" argument against nominalism. *Australasian Journal of Philosophy* 77(3): 325–334.
35. Pendlebury, M. 1986. Facts as truth-makers. *Monist* 69, 177–188.
36. Pendlebury, M. 2010. Facts and truth-making. *Topoi* 29, 137–145.
37. Prakken, H., and G. Vreeswijk. 2002. Logics for defeasible argumentation. *Handbook of Philosophical Logic* 4: 219–318.
38. Richard, M. 2008. *When truth gives out*. Oxford: Oxford University Press.

39. Rodríguez-Pereyra, G. 2006. Truth-making, entailment, and the conjunction thesis. *Mind* 115: 957–982.
40. Rosen, G. 2009. Metaphysical dependence: Grounding and reduction. In *Metaphysical grounding: Understanding the structure of reality*, ed. F. Correia and B. Schnieder, 109–136. Cambridge University Press. Cambridge
41. Schiffer, S. 2012. Truth-bearers: What are they good for? In *Current issues in theoretical philosophy: Prospects for meaning*, vol. 3, ed. R. Schantz, 531–552. Walter de Gruyter. Berlin
42. Schnieder, B. 2006. Truth-making without truth-makers. *Synthese* 152: 21–47.
43. Schulte, P. 2010. Truth-makers: a tale of two explanatory projects. *Synthese* 181(3): 413–431.
44. Soames, S. 2010. *What is meaning*? Princeton: Princeton University Press.
45. Tałasiewicz, M., J. Odrowaz-Sypniewska, W. Wciorka, and P. Wilkin. 2012. Do we need a new theory of truthmaking? Some comments on Disjunction Thesis, Conjunction Thesis, Entailment Principle and explanation. *Philosophical Studies*, doi:10.1007/s11098-012-9964-x.
46. Van Fraassen, B. 1977. The pragmatics of explanation. *American Philosophical Quarterly* 14(2): 143–150.

Chapter 16
Absolute Generality and Semantic Pessimism

J.P. Studd

Abstract Semantic pessimism has sometimes been used to argue in favour of absolutism about quantifiers, the view, to a first approximation, that quantifiers in natural or artificial languages sometimes range over a domain comprising *absolutely* everything. Williamson argues that, by her lights, the relativist who opposes this view cannot state the semantics she wishes to attach to quantifiers in a suitable metalanguage. This chapter argues that this claim is sensitive to both the version of relativism in question and the sort of semantic theory in play. Restrictionist and expansionist variants of relativism should be distinguished. While restrictionists face the difficulties Williamson presses in stating the truth-conditions she wishes to ascribe to quantified sentences in the familiar quasi-homophonic style associated with Tarski and Davidson, the expansionist does not. In fact, not only does the expansionist fare no worse than the absolutist with respect to semantic optimism, for certain styles of semantic theory, she fares better. In the case of the extensional semantics of so called 'generalised quantifiers', famously applied to natural language by Barwise and Cooper, it is argued that expansionists enjoy optimism and absolutists face a significant measure of pessimism.

16.1 Introduction

16.1.1 Absolute Generality

Absolutism about quantifiers, to a first approximation, is the view that quantifiers in natural or artificial languages sometimes range over a domain comprising *absolutely* everything. *Relativism about quantifiers* is the opposing view.

Absolutists may accept, of course, that quantifiers are *sometimes* restricted. These restrictions may be explicit in the syntax of the quantifier, as 'every donkey' or 'most universities in the Russell group' are respectively restricted to range only over the domain of donkeys and the domain of universities in the Russell group. Or, many

J.P. Studd (✉)
Lady Margaret Hall, University of Oxford, Norham Gardens, Oxford, OX2 6QA, UK
e-mail: james.studd@philosophy.ox.ac.uk

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_16

absolutists contend, even when no non-vacuous restriction is explicit in the syntax, as with syntactically unrestricted quantifiers, such as ‘nothing’ or ‘most objects’, the quantifier’s domain may nevertheless be subject to restrictions supplied by the context of utterance: a librarian’s disgruntled utterance, ‘*nothing* was returned on time last term,’ says only that *no book due back last term* was returned on time last term.[1] The operation of any sort of quantifier domain restriction is perfectly consonant with absolutism provided it can sometimes be lifted. The absolutist need only claim that some languages contain quantifiers without syntactic restrictions which in some contexts range over a domain comprising *absolutely* everything. We shall suppose he adds—as he typically does—that English is such a language and the context of his utterance attempting to expound his view is such a context. (Indeed he must add this if the exposition of his view is to achieve its required generality.)

Attempting to give a neutral characterisation of the absolutism/relativism dispute has proved notoriously difficult and it is not our concern in this chapter to resolve this issue.[2] Two preliminary clarifications are worth making all the same. The first concerns ‘domain’-talk. Absolutism should not be mistaken for the mathematically revisionary view that some *set* is a domain that comprises everything. For, as is widely recognised,[3] this view is refuted by standard set theory when axiomatised to account for non-sets or ‘urelements’. Zermelo-Fraenkel set theory with choice and urelements (ZFCU) has as a theorem the formalisation of the claim that no set has everything as an element. This claim, however, poses no threat to the view absolutists actually defend. Absolutists roundly reject what Richard Cartwright dubs the All-in-One principle.

All-in-One To quantify over some objects presupposes that those objects constitute a set or set-like collection that has those objects as its elements.

As Cartwright observes, to quantify over the biscuits in the tin, say, does not require that there be a set-domain that has those biscuits as its elements. Whether or not there in fact is such a set, the needs of this restricted quantification are met simply by there being those biscuits, severally. No further object is required to function as the domain [7, pp. 7–8]. By the same token, to quantify over *absolutely* all objects whatsoever does not require—*per impossibile*—that there be a set- or set-like domain that has *absolutely* everything as an element. The needs—at least, the ontological needs—of *absolutely* general quantification are met simply by there being some things that severally comprise *absolutely* everything.[4]

This important point, however, need not lead to an outright ban on useful ‘domain’-talk provided we are careful not to take it to carry a commitment to set- or

[1] See, for instance, Williamson [48, sec. II]. Kent Bach [1] provides an opposing view; see Stanley and Szabó [44] for discussion.

[2] Lewis [21], McGee [28] and Williamson [48, §V] press the concern that relativism cannot be coherently stated. Glanzberg [14], Fine [12] and Lavine [19] reply on behalf of the relativist.

[3] See, for instance, Cartwright [7, p. 7], Williamson [48, p. 425], and Rayo and Uzquiano [39, p. 6].

[4] Compare Boolos [5, pp. 223–4].

set-like domains. Instead, again following Cartwright, 'domain'-talk may be taken as a harmless façon de parler, elliptical for a suitable plural paraphrase. For example, 'the domain of the quantifier 'every set' comprises all sets' may be understood as elliptical for its plural paraphrase, 'the quantifier 'every set' ranges over the sets', which carries no commitment to a set-like collection of all sets[5] [7, p. 3]. In effect, Cartwright adopts a No Domain Theory of Domains analogous to Russell's No Class Theory of Classes.[6] With this point in mind, we shall continue to indulge in 'domain'-talk for the remainder of the chapter, leaving absolutists to paraphrase as we go along.

The second clarification concerns the use of '*absolutely*'. With the No Domain Theory of Domains in place, the absolutist may characterise his view as the claim that quantifiers sometimes range over some things that comprise *absolutely* everything. Relativism, however, should not be mistaken for the logically revisionary view that no things comprise everything. This view is refuted by plural first-order logic (PFO), which has as a theorem the formalisation of the claim that some things comprise everything.[7] Again, this claim poses no threat to the view relativists actually defend. From the relativist's point of view, it amounts to the trivial truth that, in the context of the utterance above, some members of the domain of the English quantifiers 'some things' and 'everything' comprise every member of that same domain. The same is true of *any* domain D, no matter how limited. To state that something lies outside a limited domain D we must quantify over a wider domain. An analogous point applies to the domain of the English quantifiers in the sentence above. To state its limitations, if any, the relativist must induce a shift to a wider domain and then—from this more liberal perspective—claim that something lies outside the initial domain. The absolutist may claim that the role of '*absolutely*' in '*absolutely* everything' is to indicate that all such shifts have been made, so that his quantifier ranges over the most liberal domain. This option is not, of course, open to the relativist. Instead, the relativist may attempt to paraphrase such claims using schemas, or other non-quantificational means of generalisation.[8]

The upshot of these two clarifications is that we should not expect the absolutism/relativism debate to be resolved by elementary considerations in set theory or logic. A number of other lines of argument, however, have been forthcoming.

[5]To avoid such commitment, the absolutist needs to maintain, contrary to Quine [34, 35], that plural reference to and quantification over objects in English is not disguised singular reference and quantification. We shall make this assumption throughout. See Boolos [3, 4] for an influential case in favour of treating plural quantification in plural terms.

[6]See, for instance, Russell [41].

[7]PFO is presented in Linnebo [24].

[8]Parsons appeals to what he calls the 'systematic ambiguity' of certain sentences to achieve such generality [31, 32]. See also Glanzberg [14]. Lavine [19] develops a relativist-friendly account of schematic generality. A different approach employed by Fine [12] is to introduce suitable modal operators, allowing us to recapture absolute generality by embedding our quantifiers within them. The resulting view falls somewhere in between absolutism and relativism, as traditionally conceived, and we set it aside here.

Relativists argue that absolutism fails to do justice to the open-ended nature of concepts such as *set* or *interpretation*,[9] or commits the absolutist to objectionable views in metaontology.[10] Absolutists in turn argue that the limits the relativist posits on quantification leave her unable to adequately capture the requisite generality for systematic philosophical, mathematical or scientific theorising of certain kinds.[11] Our concern in this chapter will be with an instance of this absolutist objection against relativism.

16.1.2 Semantic Pessimism

One field that absolutists have argued requires quantification over a domain comprising *absolutely* everything is the semantics of quantifiers themselves. Timothy Williamson [48, secs. II, VII–VIII] argues that absolutists can successfully state the semantics for quantifiers that range over a domain which he claims to comprise *absolutely* everything; but, on modest seeming assumptions, he argues that the relativist cannot do the same for quantifiers ranging over domains which she claims to be limited. Consequently, while the absolutist is well-equipped to theorise in this respect about the world as he sees it, the relativist struggles, given the limitations she posits, to capture the semantic behaviour of quantifiers as they behave according to her theory. Williamson's argument supports the following theses.

Absolutist optimism By his lights, the absolutist can state the semantics he wishes to ascribe to quantifiers in a suitable metalanguage.
Relativist pessimism By her lights, the relativist cannot state the semantics she wishes to ascribe to quantifiers in a suitable metalanguage.

Williamson concludes that 'if we can adequately state the semantics of our own language in a suitable meta-language, then generality-absolutism is true' [48, p. 449].

This result, if it could be sustained, would not be an immediately decisive consideration in favour of absolutism against relativism. It remains to be seen that semantic pessimism is untenable. Perhaps certain aspects of the world are too chaotic or complex to be brought within the range of systematic theories that are simple enough for us to comprehend.[12] Nonetheless, especially in the case of natural

[9] See, for instance, Russell [40], Dummett [9, chs. 15–6; 10, ch. 24] and, more recently, Glanzberg [14, 15], Fine [12], Hellman [16], Shapiro and Wright [43]. Boolos [5], Cartwright [7] and Williamson [48] respond on behalf of the absolutist.

[10] See Hellman [16] and Parsons [32].

[11] See Williamson [48, secs. VI–VIII]. Glanzberg [14], Fine [12], Parsons [32] and Lavine [19] respond on behalf of the relativist.

[12] Compare Williamson [48, p. 449] and Linnebo [23, p. 150].

languages, which we have an independent reason to take to be intelligible, optimism about a systematic account seems by far the more appealing option.

The aim of this chapter, however, is to show that the relationship between absolutism/relativism and semantic optimism/pessimism does not straightforwardly favour absolutism in the way Williamson claims. Instead, whether absolutists and relativists are able to successfully give semantic theories for quantifiers is sensitive to two factors: the version of relativism under consideration and the sort of semantic theory in play. The plan is as follows. The next section reviews Williamson's arguments for the claims that absolutists enjoy semantic optimism (Sect. 16.2.1) and that relativists face semantic pessimism (Sect. 16.2.2), in the case of quasi-homophonic truth theories of the sort associated with Tarski [46] and Davidson [8]. After distinguishing between restrictionist and expansionist variants of relativism (Sect. 16.3.1), it is then argued that only the former is under threat from Williamson's argument (Sect. 16.3.2). In fact, not only does the expansionist fare no worse than the absolutist with respect to semantic optimism, for certain styles of semantic theory, she fares better. Section 16.4 briefly reviews a second style of semantic theory, Barwise and Cooper's [2] influential application of the extensional semantics for quantifiers developed by Mostowski [30] to natural language quantifiers (Sect. 16.4.1), and argues that expansionists enjoy optimism and absolutists face pessimism when it comes to stating semantic theories in this style in the standard way. Section 16.5 then evaluates the prospects (Sect. 16.5.1) and costs (Sect. 16.5.2) of an absolutist-friendly recasting of such semantics in an artificial metalanguage with superplural expressive resources.

16.2 Tarski-Davidson

16.2.1 Initial Absolutist Optimism

First-order quantification over a given universe is ill-suited for theorising about every possible interpretation of our expressions over that same universe. Following Agustín Rayo, let us call a model theory for a language *strictly adequate* if every possible extension of an expression is assigned to that expression in some model. By the absolutist's lights, we cannot give a strictly adequate model theory for a first-order language whose quantifiers range over *absolutely* everything in a first-order metalanguage. Of course the intended extensions of some predicates, such as 'thing' or 'set', comprise too many objects to form a set. But even without assuming that extensions are set-extensions, a suitable version of Cantor's theorem shows that there are more possible predicate-extensions over a universe M than there are objects in the universe M. As Rayo shows, this theorem can be precisely stated when predicate-extensions are encoded plurally as the things to which the predicate is taken to apply. In order, therefore, to give a strictly adequate model theory for a first-order language quantifying over *absolutely* everything, we must ascend to a metalanguage with plural (or higher-order) expressive resources, which is able to quantify over every predicate-extension thus encoded. Analogous considerations force further ascent to give strictly adequate model theories for languages with plural predicates, and so on. The conclusion Rayo draws is that if the absolutist is to avoid the pessimistic conclusion that it is impossible to give a strictly adequate model-theory for a language countenanced as legitimate—a conclusion we might label model-theoretic pessimism—he must endorse an open-ended hierarchy of languages of ever ascending logical type [37, esp. secs. 9.6–9.7].

Model-theoretic pessimism, however, should be distinguished from pessimism about the prospects for giving the semantics for a particular language, where we are concerned with just a single interpretation of the language's expressions. In this chapter, we shall be primarily concerned with the latter form of semantic pessimism. And here some initial optimism is not hard for the absolutist to come by, without the need to countenance an open-ended hierarchy of plural (or higher-order) resources.

Consider the case of an interpreted first-order language $\mathcal{L}_{\text{A}}$, whose universal quantifier $\forall_{\text{A}}$ is claimed by the absolutist to range over *absolutely* everything.[13] Such a language provides a simple model of the semantics the absolutist attaches to syntactically unrestricted quantifiers such as 'everything' when no contextual restrictions are operative.

As Williamson observes, the absolutist can formulate the truth-conditions he attaches to sentences of the form $\forall_{\text{A}} v\phi$ in the familiar quasi-homophonic style associated with Tarski [46] and Davidson [8]. He may give the following satisfaction clause for the quantifier $\forall_{\text{A}}$:[14]

(S-$\forall_{\text{A}}$) $\forall_{\text{A}} v\phi$ is true under assignment σ iff everything a is such that ϕ is true under $\sigma[v/a]$.

Of course, if this is to ensure that $\forall_{\text{A}}$ expresses universal quantification over *absolutely* everything, as the absolutist intends, then the metalanguage's quantifier 'everything' must likewise range over a domain that comprises *absolutely* everything [48, p. 418].

Consequently, absolutists and relativists will differ on the success of (S-$\forall_{\text{A}}$). By the relativist's lights, the metalanguage's quantifier 'everything', like any other, only ranges over a limited domain and, as such, the stipulated satisfaction clause renders $\forall_{\text{A}}$ similarly limited. But the absolutist admits no such obstacle. In his view—at least in unrestricted contexts—the English quantifier 'everything' deployed in the metalanguage expresses universal quantification over *absolutely* everything, and may thus succeed in stating the semantics he attaches to $\forall_{\text{A}}$. The absolutist attains the following amount of semantic optimism for Tarski-Davidson style semantic theorising.

[13]We adopt the logician's convention of omitting quotes from expression from formal languages.

[14]Assignments may be treated in the standard way as (set-)functions from variables to objects. As usual, the assignment $\sigma[v/a]$ agrees with σ on every variable other than v and maps v to a.

Absolutist optimism By his lights, the absolutist *can* state the truth-conditions he wishes to ascribe to the quantifier $\forall_{\mathrm{A}}$ in a metalanguage quantifying over the same domain as the object language.

16.2.2 Williamson's Case for Relativist Pessimism

Williamson [48, sec. VII] argues that relativists do not enjoy parallel semantic optimism: the relativist is unable, by her lights, to state the semantics she wishes to ascribe to quantifiers in the Tarski-Davidson style.

Williamson focuses on the case of an interpreted first-order language 'fit simply for the expression of context-bound generality, a language of the sort one might expect to be innocuous from the perspective of generality-relativism' [48, p. 445]. Let us call this language $\mathcal{L}_C$. In each context c, the universal quantifier $\forall_{\mathrm{C}}$ of $\mathcal{L}_C$ expresses universal quantification over a contextually-determined domain D^c. Williamson gives the natural satisfaction clause for $\forall_{\mathrm{C}}$ [48, p. 444]:

(S-$\forall_{\mathrm{C}}$) For each context c, $\forall_{\mathrm{C}} v\phi$ is true in c under assignment σ iff every object a that is a member of the domain D^c is such that ϕ is true under $\sigma[v/a]$.

From the absolutist's perspective, even if in some contexts the quantifier $\forall_{\mathrm{C}}$ ranges over *absolutely* everything, (S-$\forall_{\mathrm{C}}$) succeeds in stating the truth-conditions he ascribes to sentences of the form $\forall_{\mathrm{C}} v\phi$ in every context provided the metalanguage's quantifier 'every object' also ranges over *absolutely* everything.

The relativist, of course, will insist that the object language's quantifier is always limited. Williamson argues, however, that even when we suppose the range of $\forall_{\mathrm{C}}$ to always be thus limited, the relativist is unable to successfully capture its semantics with a Tarski-Davidson style satisfaction clause. The essentials of his case for pessimism may be summarised in three claims. Suppose the relativist attempts to specify the intended truth-conditions for sentences of the form $\forall_{\mathrm{C}} v\phi$ by uttering (S-$\forall_{\mathrm{C}}$) in a theoretic context $c^\star$ in which the metalanguage's quantifier 'every object' ranges over $D^\star$. Observe first that the relativist's utterance only captures the intended semantics for the quantifier if the theoretical context's domain $D^\star$ is at least as inclusive as every domain D^c that $\forall_{\mathrm{C}}$ ranges over.

(P1) The utterance of (S-$\forall_{\mathrm{C}}$) in $c^\star$ specifies the intended truth-conditions for sentences of the form $\forall_{\mathrm{C}} v\phi$ only if each one of the things in the domain D^c of $\forall_{\mathrm{C}}$ in any context c is a member of $D^\star$.

This is because the limitations on the metalanguage's quantifier posited by the relativist have the effect of limiting the object language's quantifier $\forall_{\mathrm{C}}$ twice over in each context c. The effect of the satisfaction clause (S-$\forall_{\mathrm{C}}$) as uttered in $c^\star$ is to specify that, for each context c, $\forall_{\mathrm{C}} v\phi$ is true in c under an assignment σ just in case every *member a of* $D^\star$ that is a member of the domain D^c is such that ϕ is true under $\sigma[v/a]$—just in case, in other words, every member of the intersection of the metalanguage- and contextual-domain, $D^\star \cap D^c$, is such that ϕ is true under

$\sigma[v/a]$. The net effect is to specify that, in each context c, $\forall_C$ expresses universal quantification over $D^\star \cap D^c$. Consequently, the relativist's utterance ascribes $\forall_C$ its intended interpretation of universal quantification over D^c in each context c only if $D^\star \cap D^c = D^c$ for each context c, and thus each one of the things in the domain D^c of any context c is a member of $D^\star$.[15] (Compare [48, p. 445].)

Second, Williamson argues that the relativist has reason to allow that while each domain D^c is limited, everything is in the domain of some context or other:

(P2) Everything is in the domain D^c of at least one context c.

Williamson argues that the relativist needs to accept this in order to avoid what he calls 'semantic pariahs' for $\mathcal{L}_C$: objects that fall outside the domain of that language's quantifier in every context. In his view, 'it is highly implausible to think that there are such semantic pariahs for English; natural language quantification seems too promiscuous for that' [48, pp. 445–6]. Moreover, since the relativist has no reason to take $\mathcal{L}_C$ to be more restricted than she takes natural languages to be, Williamson concludes that it would be equally implausible to posit such 'pariahs' for languages like $\mathcal{L}_C$.

Third, and last, the relativist does not, of course, accept that the domain $D^\star$ has *absolutely* everything as a member. Rather, she accepts the following:

(P3) Not everything is in $D^\star$.

Here, care must be taken with the context. The utterance of (P3) only succeeds in its intent if the context of utterance—$c^{\star\star}$, let's say—is more liberal than that of the relativist's original theoretical context $c^\star$ in which she laid down the satisfaction clause (S-$\forall_C$). Uttered in $c^\star$, (P3) expresses the trivial falsehood that not every member of $D^\star$ is in $D^\star$. Provided that we are careful with the choice of $c^{\star\star}$, however, there seems to be good reason for the relativist to accept what is expressed by each premiss (P1), (P2) and (P3) in $c^{\star\star}$. The conclusion (C) follows from the premisses (P1)–(P3):

(C) The utterance of (S-$\forall_C$) in $c^\star$ fails to specify the intended truth-conditions for sentences of the form $\forall_C v\phi$.

Once this has been made clear, there appears to be a compelling reason for the relativist to concede that she is committed to semantic pessimism. From the relativist's perspective, the natural attempt to state the truth-conditions for quantified sentences in a relativist-friendly language whose quantifiers always range over a limited domain fails.

[15]The analogues of elementary set-theoretic operations like intersection and union are readily accommodated within the No Domain Theory of Domains. $D^\star \cap D^c$ may be treated as a plural term denoting the things that are both one of the members of $D^\star$ and one of the members of D^c.

16.3 Restrictionism vs. Expansionism

16.3.1 Domains and Universes

To assess the general effectiveness of this style of objection against relativism, it is helpful to distinguish two variants of the view. *Restrictionism* and *expansionism* both oppose absolutism but for different reasons: the restrictionist claims—in a sense to be elucidated—that the domains of our quantifiers are always subject to restriction, the expansionist claims that the domains of our quantifiers are always open to expansion.

The distinction is due to Kit Fine, who distinguishes two ways in which we might enlarge a quantifier's domain. The first is through *de-restriction*. On this model 'the interpretation of the quantifier is given by something like a predicate or property which serves to restrict its range'. We increase the domain by relaxing the restricting condition [12, p. 35]. For example, we might de-restrict the domain of 'every bottle' by relaxing the restricting predicate 'bottle' to apply not just to bottles but also to other things made of glass, say. Notice however that de-restricting the quantifier in this way to shift from the domain comprising every bottle to the new, wider domain of glass objects relies on our having some sort of grasp on a third totality encompassing both domains. It relies on an understanding of the determiner 'every', which when combined with a universally-applicable predicate like 'thing' and subject to no other restrictions ranges over the encompassing totality. The second means to enlarge domains does not rely on such an understanding. Rather, on the basis of understanding quantification over an initial domain, reinterpretation by *expansion* permits us to come to understand quantification over a wider domain, without presupposing any sort of grasp of a third totality encompassing both. As Fine puts it, 'We might say that the new domain is understood from 'above' under . . . the restrictionist account, insofar as it is understood as the restriction of a possibly broader domain, but that it is understood from 'below' under the expansionist account, in that it is understood as the expansion of a possibly narrower domain' [12, p. 38].[16]

To further elucidate this distinction—at least, as we shall understand it here[17]—let us distinguish *domains* from *universes*. A *domain* (as we use the term) is tied to a quantifier, relative to a specific context and interpretation. We have a fairly robust grasp of this notion. To give a circular elucidation: the domain of a quantifier in a context comprises the objects it ranges over (given its interpretation). For example the domain of the English quantifier 'every bottle' comprises all bottles

[16]Fine makes this distinction in the context of defending his third-parameter version of relativism, procedural postulationism, mentioned in n. 23, but we should separate the two. The distinction between restriction and expansion has wider significance in the debate about absolute generality.

[17]Although Fine does not gloss the distinction in these terms, it seems to provide a natural regimentation of the view he elucidates. Our primary concern, however, remains semantic pessimism rather than Fine exegesis.

(when subject to no further contextual restrictions).[18] A *universe* is tied to a whole language, the interpretation of an entire lexicon, *not* relative to a specific context. The universe of a language encompasses every object in the domain of any quantifier interpreted in the language, in any context, together with any object that is the extension of a singular term or is a member of the extension of a predicate, and so on. The universe of the language to which the quantifier 'every bottle' belongs, the universe of a version of English, includes in addition to bottles, everything else that we can refer to or quantify over in this language in any context. Consequently, relative to a fixed interpretation, the domain of a quantifier is always a subdomain of the universe.[19]

In order to succeed in quantifying over *absolutely* everything, the absolutist needs to maintain that both an intra- and inter-language barrier may be overcome: first, in some contexts, some quantifiers—including, he claims syntactically unrestricted quantifiers such as 'everything'—surpass the *intra*-language barrier by ranging unrestrictedly over the entire universe of the language they belong to; second, some languages—including, he claims, the present version of English—transcend the *inter*-language barrier by having a universe as inclusive as the universe of any other language. His claim that, in a context c, under an interpretation I, the domain D^c of a quantifier comprises *absolutely* everything may be factored into two claims:

(A1) The domain D^c is *unrestricted*: the domain of the quantifier in the context contains *absolutely* every member of the universe M of the interpretation I.

(A2) The universe M is *inexpandable*: the universe M of the interpretation I contains *absolutely* every member of *absolutely* every universe.

Let us distinguish the two types of relativist according to which claim they oppose. The *restrictionist*—or, to give her her full title, the anti-expansionist restrictionist—posits only the intra-language barrier. On this view, some languages have inexpandable universes but the domain of each quantifier belonging to such a language is always subject to restriction: its domain constitutes a proper subdomain of this universe. The domain may always be widened by relaxing the restriction but this widening occurs within the universe of the language that the speaker has some grasp on through her understanding of the language. The second version of relativism, (anti-restrictionist) *expansionism*, posits only the inter-language barrier. On this view, some quantifiers sometimes range unrestrictedly over a domain that encompasses the entire universe of the language to which they belong but this universe is always open to expansion: the universe constitutes a proper subuniverse

[18] A non-circular elucidation follows in Sect. 16.4.1.

[19] We define *subdomain* (also 'subuniverse', etc.) as the obvious analogue of subset: D is a *subdomain* of D' (in symbols: $D \subseteq D'$) if every member of D is a member of D'. A subdomain D of D' is said to be a *proper subdomain* of D' ($D \subset D'$) if, moreover, D' is not a subdomain of D.

of the universes of more inclusive languages. Speakers may come to understand more liberal languages without having an antecedent grasp of their universe.[20]

Michael Glanzberg [14, 15] defends a contextual version of relativism. On his view, reflection on the set-theoretic and semantic paradoxes always permits us to shift to a more inclusive context; iterating such shifts we can come to ever wider domains: $D_0 \subset D_1 \subset \cdots$. Notice, however, that a mere shift of context can never take us outside the universe of the language. The universe of a language, as we have characterised it, contains the objects in the domain of any quantifier that belongs to the language in any context. As such, as we have drawn the distinction between restrictionism and expansionism, this view counts as a version of restrictionism.[21] Mere shifts in context cannot take us outside the universe M of the language but instead lead to ever wider domains properly contained within the universe: $D_0 \subset D_1 \subset \cdots \subset M$. On Glanzberg's view, the shifts in domain brought about by reflection on the paradoxes are analogous to shifts in the referent of an indexical like 'I' brought about by shifts of speaker [15, pp. 50, 53]. In contrast, expansionists may take shifts in the universe to be more akin to the shift in the referent of 'Madagascar' brought about by speakers coming to use this name in a new way.[22] Shifts in universe results from shifts of interpretation rather than mere shifts of context.[23]

According to expansionism, in some contexts, syntactically unrestricted quantifiers like 'everything' may range over the entire universe M of a version of English. However, the expansionist claims that we can always come to speak a more liberal version of English, in which 'everything' expresses universal quantification over a wider domain. This shift requires us to liberalise the meaning of our expressions, to re-interpret them over a universe M' that is more inclusive than the initial universe M without an antecedent understanding of any language interpreted over M'. Expansionism thus incurs a substantial explanatory burden: advocates of this view need to give an account of how such expansion operates. Here however we shall restrict our attention to assessing how these two versions of relativism fare with respect to semantic pessimism.

---

[20]More extreme versions of relativism are also possible, according to which both barriers are imposed: domains are always restricted and universes are always expandable.

[21]We assume here that Glanzberg does not go in for the more extreme version of relativism mentioned in n. 20. Note that there is a sense in which domains are expanded on this account—shifts in context lead us to wider domains—and Glanzberg applies the label 'expansionism' to his view [15, n. 5]. Terminological issues aside, however, what matters in the context of semantic pessimism is that these domains are always proper subdomains of the universe of the entire language.

[22]We borrow Evans and Altham's [11] famous example of reference shift.

[23]Fine [12, sec. 2.6] outlines what seems to be a third option according to which shifts in universe result from a shift in 'ontology' distinguished as a third parameter, distinct from both shifts in the circumstances and shifts in semantic content.

16.3.2 Relativist Pessimism Revisited

With this distinction in hand, let us return to semantic pessimism. Recall that Williamson argues that the relativist is unable to give the semantics she wishes to ascribe to the context-sensitive quantifier $\forall_C$ by uttering the satisfaction clause (S-$\forall_C$) in a suitable theoretical context $c^\star$. The thrust of the argument outlined in Sect. 16.2.2 is that from the perspective of a suitable context $c^{\star\star}$ the relativist should accept what is expressed by each of (P1), (P2) and (P3); and consequently, she should also accept the pessimistic conclusion expressed by (C), which follows from them. Making the relativity to the domain $D^{\star\star}$ of the context $c^{\star\star}$ explicit, the premisses and conclusion are as follows:

(P1)$^{c^{\star\star}}$ The utterance of (S-$\forall_C$) in $c^\star$ specifies the intended truth-conditions for sentences of the form $\forall_C v\phi$ only if every one of the members of $D^{\star\star}$ in the domain D^c of $\forall_C$ in any context c is a member of $D^\star$.
(P2)$^{c^{\star\star}}$ Every member of $D^{\star\star}$ is in the domain D^c of at least one context c.
(P3)$^{c^{\star\star}}$ Not every member of $D^{\star\star}$ is in $D^\star$.
(C)$^{c^{\star\star}}$ The utterance of (S-$\forall_C$) in $c^\star$ fails to specify the intended truth-conditions for restrictedly quantified sentences of the form $\forall_C v\phi$.

The argument is valid but restrictionists and expansionists will differ on its soundness. On the restrictionist's view, since $\mathcal{L}_C$ is intended to supply a model of quantifier domain restriction in natural language, we may reasonably make two assumptions. First, we may assume that $\mathcal{L}_C$ is interpreted over the inexpandable universe just as she takes versions of English to be. Second, we may assume that every legitimate domain of quantification available in the metalanguage is also available for $\forall_C$ to range over in the object language, in some context; in particular, there are object language contexts c_0 and c_1 whose domains are the same as the domain of the metalanguage's quantifier in $c^\star$ and $c^{\star\star}$: $D^{c_0} = D^\star$ and $D^{c_1} = D^{\star\star}$.

Given these two assumptions, the restrictionist has little option but to accept the soundness of Williamson's argument. Both sorts of relativist should accept (P1)$^{c^{\star\star}}$ for the reasons outlined in Sect. 16.2.2. If the domain $D^\star$ of the metalanguage's quantifier lacks a member of some contextually-determined domain D^c, the satisfaction clause gives the wrong truth-conditions for sentences of the form $\forall_C v\phi$ in context c. (P2)$^{c^{\star\star}}$ holds in virtue of our second assumption. Every member of $D^{\star\star}$ is a member of D^{c_1}, and thus a member of the domain D^c of $\forall_C$ in some context. Finally (P3)$^{c^{\star\star}}$ holds in virtue of the restriction on the metalanguage's quantifier's domain $D^\star$. Granted that the shift from $c^\star$ to $c^{\star\star}$ results in a wider domain $D^{\star\star}$—as the restrictionist contends is possible—this premiss holds.

The expansionist is in a different position. In order for $\mathcal{L}_C$ to provide a model of natural language quantification, this sort of relativist is under no pressure to claim that the object language ranges over the same universe as the version of English deployed as the metalanguage. For she does not claim that all versions of English are interpreted over the same universe. Instead she characteristically claims that the universes of natural languages are always open to expansion. However, for the sake of simplicity, we may assume—as the expansionist allows is possible—that the universe M of $\mathcal{L}_C$ is also the universe of the version of English deployed as the metalanguage, and that the domain $D^\star$ of the theoretical context $c^\star$ in which (S-$\forall_C$) is laid down encompasses the whole of this universe.

Given these assumptions, the expansionist should accept $(P1)^{c^{\star\star}}$ for the same reasons as before. She should also accept $(P3)^{c^{\star\star}}$ provided we are careful with the choice of $c^{\star\star}$. Crucially, however, in this case, the shift from $c^\star$ to $c^{\star\star}$ cannot merely be a shift of context but must also involve a shift of interpretation. For we have supposed that the domain $D^\star$ of the metalanguage's quantifier in $c^\star$ is the entire universe M of the version of English in which (S-$\forall_C$) is laid down. The expansionist will only accept what is expressed by the utterance of (P3) in $c^{\star\star}$ if $D^{\star\star}$ is wider than $D^\star$ (i.e. M). Consequently, this utterance must be made in a more liberal version of English to allow $D^{\star\star}$ to surpass the universe M of the version of English in which (S-$\forall_C$) was laid down. Given this shift of language, the expansionist is under no pressure to accept the truth of $(P2)^{c^{\star\star}}$. This premiss is true only if each member of $D^{\star\star}$ is a member of a domain of the object language's quantifier, and thus a member of the object language's universe M. But the expansionist denies this. The domain $D^{\star\star}$ of the metalanguage's quantifier in the new, more liberal version of English is wider than the universe M of the initial metalanguage. The expansionist is not under the same pressure as the restrictionist to concede the soundness of Williamson's argument.

In rejecting $(P2)^{c^{\star\star}}$ does the expansionist commit herself to 'semantic pariahs'? She does, of course, allow that the universe of the new version of English in which Williamson's argument is stated contains objects not in the universe of the version of English in which the satisfaction clause (S-$\forall_C$) was set out. Such objects are beyond the reach of quantification in the initial version of English but not beyond the reach of natural language quantification in general. After all, they are quantified over in the new version of English. Any 'pariah' status is merely temporary. Might the absolutist nonetheless press the objection and contend that it is highly implausible that a particular natural language admits such temporary 'pariahs'? Such a claim comes too close to a flat out denial of an aspect of her view to be dialectically effective against the expansionist. The claim that we can effect semantic change so as to expand the universe of a natural language and thereby come to quantify over objects not quantified over in the initial language is at the centre of expansionism. And while, as we noted above, the proponent of this view owes us an account of just how such expansion is achieved, it is not obvious that this sort of semantic change is impossible.

The distinction between restrictionism and expansionism made in the previous section consequently leads to a real difference when it comes to semantic pessimism with respect to languages like $\mathcal{L}_C$ intended to provide a simple model of quantifier domain restriction in natural language. For the restrictionist, there is no metalanguage context $c^\star$ that is sufficiently unrestricted to enable her to state a Tarski-Davidson style satisfaction clause for $\forall_C$, giving the semantics she wishes to ascribe to the quantifier in each context. For the expansionist—just like in the case of the absolutist—a metalanguage context $c^\star$ whose domain $D^\star$ encompasses the entire object language's universe is liberal enough to enable her to give just such a clause.

Restrictionist pessimism By her lights, the restrictionist cannot state the truth-conditions she wishes to ascribe to sentences of the form $\forall_C v\phi$, in each context, by quantifying over a proper subdomain of the object language's universe in the metalanguage.

Expansionist optimism By her lights, the expansionist can state the truth-conditions she wishes to ascribe to sentences of the form $\forall_C v\phi$, in each context, by quantifying over the whole of the object language's universe in the metalanguage.

The restrictionist, of course, can still give the intended semantics for *portions* of the object language, stating the truth-conditions for every context whose domain is contained within the available theoretical domain. She can moreover schematically indicate the form the satisfaction clause will take in wider contexts. However, the pessimistic conclusion remains that, contrary to the efforts of semanticists who attempt to give semantic theories for *entire* languages, no theoretical context is available that is liberal enough to state the semantics for the whole language.

There is another way in which the restrictionist may attempt to ameliorate this pessimistic conclusion. Her facing semantic pessimism in the case of artificial languages like $\mathcal{L}_C$ leaves open the question of whether she also faces a more damaging form of pessimism, semantic pessimism in the case of natural languages like English, as these function according to restrictionism. And while Williamson has chosen $\mathcal{L}_C$ to model relativist-friendly context-dependent quantification, there are some well-known reasons to think that quantifier domain restriction in natural language functions differently to the quantifier domain restriction stipulated in the semantics for $\mathcal{L}_C$.

Consider the following example from Stanley and Szabó [44, p. 249]. Suppose that as the boat leaves the harbour, all the sailors stand on deck and wave to all the sailors on the shore who wave back. It seems that in such a context, an utterance of:

(1) Every sailor waved to every sailor

says that every sailor *on the boat* waved to every sailor *on the shore*.[24] Or again, suppose that everyone who came along to the party brought some bottles with them, and for some reason or other drank only what they had contributed. It seems that uttered in such a context:

(2) Everyone drank every bottle

says that everyone *who came* drank every bottle *that they brought with them*.[25] Neither of these readings can be successfully captured if, as under the stipulated

[24] See also the example attributed to Peter Ludlow by Stanley and Williamson [45].

[25] Williamson [48, p. 419] presents a similar example; compare the example presented by Stanley and Szabó [44, p. 243].

semantics for $\forall_C$, we take the context c to determine a single domain D^c over which all quantifiers in the sentence range, subject only to further restrictions explicit in their syntax.

In light of such examples, Michael Glanzberg [15] identifies two different kinds of contextual restriction in addition to the restrictions due to constituents explicit in the syntax of quantifiers. First, there are local contextual restrictions. Following Stanley and Szabó [44], Glanzberg allows that the context may impose different local restrictions on different occurrences of quantifiers occurring in a single sentence. On such an account, the context of utterance of (1) supplies tacit constituents that restrict the first occurrence of 'every sailor' to sailors on the boat and the second occurrence of 'every sailor' to sailors on the shore. In the case of (2), the context supplies a restricting constituent with a variable that is bound by the outer quantifier, restricting 'every bottle' to the bottles brought by x for each value of x that 'everyone' ranges over. Second, there is a background restriction to a single, contextually-determined background domain, operating in a manner much like the single, contextually determined domain D^c that $\forall_C$ is stipulated to range over in (S-$\forall_C$) [15, pp. 49–54].

Glanzberg contends that while it is local contextual restrictions that do the work in accounting for the truth-conditions of utterances such as (1) and (2), it is background domain relativity that is at the root of relativism. In his view, while occurrences of quantifiers such as 'everything' are syntactically unrestricted and, in some contexts, subject to no local contextual restrictions, so that they range over the entire background domain in that context, reflection on the paradoxes always permits us to shift to a more extensive background domain [15, pp. 50, 60–2].

Sophisticated versions of restrictionism like the one espoused by Glanzberg are well-placed to ensure that sentences like (1) and (2) obtain the correct truth-conditions. But these refinements do nothing to alleviate the problem of semantic pessimism for natural languages interpreted over an inexpandable universe. For we still need to state the truth-conditions for sentences containing the syntactically unrestricted quantifier 'everything' in each context c in which no local restrictions come into play. According to Glanzberg such a quantifier ranges over a background domain, which we may label D^c. The natural satisfaction clause will closely resemble (S-$\forall_C$) and face exactly the same problem. The clause will only capture the intended truth-conditions if the metalanguage's quantifier used to state these truth-conditions ranges over a domain $D^\star$ at least as inclusive as every background domain D^c. But this cannot be so. For reflection on the paradoxes can lead us to a strictly wider background domain $D^{\star\star}$. The embellished version of restrictionism faces pessimism in the case of natural languages for much the same reason that it faces pessimism in the case of the artificial language $\mathcal{L}_C$.[26]

On the other hand, absolutist and expansionist optimism about $\mathcal{L}_C$ seems to extend to optimism about Tarski-Davidson style semantics for natural language too. There is considerable room for disagreement between absolutists about the

[26]Williamson generalises his argument to sortal versions of restrictionism [48, sec. VIII].

semantics of quantifier domain restriction in natural language. Williamson does away with the second sort of contextual restriction posited by Glanzberg, dispensing with the relativisation to a background domain in the satisfaction clause for 'every'. On this view, quantification over the entire universe may be achieved by syntactically unrestricted quantifiers like 'everything' in contexts where no local contextual restrictions come into play [48, sec. II]. Kent Bach [1] goes one step further, and also does away with local contextual restrictions, preferring instead to draw on pragmatics to deal with examples like (1) and (2). On this view, syntactically unrestricted quantifiers like 'everything' range over the entire universe *in every context*. Both these options are also open to the expansionist, who can allow that such syntactically unrestricted quantifiers are sometimes or always wholly unrestricted, ranging over the entire universe. And in both cases, the availability of such unrestricted quantification permits the absolutist and expansionist to state the truth-conditions each ascribes to quantified sentences in every context.

16.4 Mostowski-Barwise-Cooper

16.4.1 Extensional Semantics

Absolutists may be rightly optimistic about their prospects for giving Tarski-Davidson style semantic theories for the semantics they ascribe to quantifiers. But not all semanticists opt for this style of semantic theory. In their seminal paper, Barwise and Cooper [2] apply the extensional semantics for so-called 'generalised quantifiers' to the semantics of quantifiers and determiners in formal and natural languages.[27] What becomes of semantic optimism and pessimism in this setting?

Let us first briefly review this approach. Start with syntax. Unlike in $\mathcal{L}_C$ where quantifiers are variable binding sentence operators, in English quantifiers are noun phrases. A quantifier consists of a determiner ('every', 'some', 'no', 'most', etc.), together with a nominal, a singular or plural noun, perhaps qualified with adjectives or relative clauses ('thing', 'sailor', 'bottles that are on the table', etc.). Quantifiers combine with verb phrases to form sentences [2, pp. 161–2].[28] Barwise and Cooper deploy a formal language $\mathcal{L}_{\text{GQ}}$ whose syntax corresponds more closely to that of natural language. $\mathcal{L}_{\text{GQ}}$ extends a first-order language by adding determiners such as EVERY, SOME, NO and MOST, together with a distinguished unary predicate THING. A determiner d combines with a unary predicate η, to form a quantifier $d(\eta)$. A quantifier $d(\eta)$ combines with a unary predicate θ to form a sentence $d(\eta)(\theta)$. Variable binding, when required, may be carried out by using λ-abstraction to form complex predicates [2, p. 168].

[27] Barwise and Cooper often use the label 'model-theoretic semantics'. We deviate from their terminology to avoid blurring the distinction between model theory and semantics.

[28] See also Lewis [20, p. 40].

Turn to semantics. The idea, very roughly, is to conceive of the meaning of a quantifier as a second-order property. For example, the meaning of 'something' is the second-order property of being an instantiated first-order property and the meaning of 'everything' is the second-order property of being a universally instantiated first-order property. 'Something is a bottle' is true since the first-order property of being a bottle has the second-order property of being instantiated; 'Everything is a bottle' is false since the first-order property of being a bottle lacks the second-order property of being universally instantiated.

This thought, in essentials, goes back to Frege,[29] and was rediscovered in its extensionally-sanitised modern form by Mostowski [30] and generalised by Lindström [22]. The standard kind of extensional interpretation of a first-order language—the sort one gives by specifying a model supplying extensions for predicates, singular terms, and so on—is generalised in a natural way to also assign extensions to quantifiers and determiners. As usual, syntactically simple singular terms, predicates and sentences are assigned extensions appropriate to their syntactic category. Let us write $|e|$ for the extension of an expression e, and use boldface $\boldsymbol{e}$ to denote extensions of that category. The extension of a singular term τ is an element $\boldsymbol{\tau}$ of the universe M; the extension of a unary predicate θ is a unary relation $\boldsymbol{\theta}$ on M (a set of elements of M); the extension of an n-ary predicate θ is an n-ary relation $\boldsymbol{\theta}$ on M (a set of n-tuples of members of M); the extension of a sentence is a truth-value. But extensions are also assigned to quantifiers such as EVERY(THING) and determiners such as EVERY. The extension of a quantifier q is a unary relation $\boldsymbol{q}$ on unary-predicate-extensions (a set of sets of members of M); the extension of a determiner d is a function $\boldsymbol{d}$ mapping predicate-extensions to quantifier-extensions. The intended extension of THING is the universe M; the intended extensions for EVERY, SOME, NO and MOST are defined as follows, for each predicate-extension $\boldsymbol{\eta}$[30]:

$$|\text{EVERY}|(\boldsymbol{\eta}) = \{\boldsymbol{\theta} \subseteq M \mid \boldsymbol{\eta} \subseteq \boldsymbol{\theta}\} \qquad |\text{SOME}|(\boldsymbol{\eta}) = \{\boldsymbol{\theta} \subseteq M \mid \boldsymbol{\eta} \cap \boldsymbol{\theta} \neq \emptyset\}$$

$$|\text{NO}|(\boldsymbol{\eta}) = \{\boldsymbol{\theta} \subseteq M \mid \boldsymbol{\eta} \cap \boldsymbol{\theta} = \emptyset\} \qquad |\text{MOST}|(\boldsymbol{\eta}) = \{\boldsymbol{\theta} \subseteq M \mid |\boldsymbol{\eta} \cap \boldsymbol{\theta}| > |\boldsymbol{\eta} - \boldsymbol{\theta}|\}$$

The extensions of complex expressions are determined according to the natural compositionality clauses. The compositionality clauses for unary predicates, quantifiers and determiners are as follows.

(C-θ) For any unary-predicate-extension $\boldsymbol{\theta}$ and singular-term-extension $\boldsymbol{\tau}$, and any unary predicate θ and singular term τ with $|\theta| = \boldsymbol{\theta}$ and $|\tau| = \boldsymbol{\tau}$: $|\theta(\tau)| = \text{T}$ if and only if $\boldsymbol{\tau} \in \boldsymbol{\theta}$.

[29]See, for instance, Frege [13].

[30]Here MOST is taken to have its weakest sense; so interpreted, MOST$(\eta)(\theta)$ says roughly that more than half of the satisfiers of η satisfy θ.

(C-q) For any quantifier-extension $\boldsymbol{q}$ and unary-predicate-extension $\boldsymbol{\theta}$, and any quantifier q and unary predicate θ with $|q| = \boldsymbol{q}$ and $|\theta| = \boldsymbol{\theta}$: $|q(\theta)| = \mathrm{T}$ if and only if $\boldsymbol{\theta} \in \boldsymbol{q}$.

(C-d) For any determiner-extension $\boldsymbol{d}$ and unary predicate-extension $\boldsymbol{\eta}$, and any determiner d and unary predicate η with $|d| = \boldsymbol{d}$ and $|\eta| = \boldsymbol{\eta}$: $|d(\eta)| = \boldsymbol{d}(\boldsymbol{\eta})$.

For example, the sentence EVERY(THING)(BOTTLE) is true just in case |BOTTLE| ∈ |EVERY(THING)|; this holds just in case |THING| ⊆ |BOTTLE|; that is, just in case every member of the universe is in |BOTTLE|. These compositionality clauses may be generalised in the natural way to predicates (and quantifiers) of greater arity. (Compare [2, sec. 2.5].)[31]

The extensional approach to the semantics of quantifiers in natural language has proved enormously fruitful. As Barwise and Cooper observe, encoding the meanings of quantifiers and determiners, in addition to those of singular terms, predicates and sentences, as extensions permits us to give the intended semantics for determiners like MOST which cannot be captured as unary quantifiers in first-order languages [2, thm. C13]. It also facilitates the mathematical investigation of the properties and relations that hold of quantifiers and determiners, and has permitted theoretically inclined semanticists to generate a wealth of putative linguistic universals to be tested in the field. To give just one example, Barwise and Cooper hypothesise that natural language quantifiers of the form corresponding to $d(\eta)$ always 'live on' the extension of the nominal predicate: this is to say that, for any predicate-extensions $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ and natural language determiner-extension $\boldsymbol{d}$:

$$\boldsymbol{\theta} \in \boldsymbol{d}(\boldsymbol{\eta}) \text{ iff } \boldsymbol{\theta} \cap \boldsymbol{\eta} \in \boldsymbol{d}(\boldsymbol{\eta}).$$

[2, sec. 4.4.] This property has come to be known as conservativity, and provides precise, empirically testable content to our inchoate pre-theoretic sense that quantifiers of the form $d(\eta)$, such as 'every bottle on the table', are restricted by the nominal η. Consider, for instance, the sentence 'every bottle on the table is empty'. On the extensional account this sentence is true if the extension of 'empty' has a certain extensional second-order property, namely the property of containing every bottle on the table (i.e. if |'empty'| ∈ |'every bottle on the table'|). The sense in which the quantifier is restricted manifests itself in the fact that we need not look at the entire extension of 'empty', examining every empty thing in the entire world, in order to determine whether the extension has the second-order property. In view of conservativity, the extension of 'empty' has the property in question, the property encoded by |'every bottle on the table'|, just in case its restriction to bottles on the

[31] The extensional approach may be naturally generalised to intensional languages. Since issues pertaining to intensionality do not concern us here, we continue to simplify by focusing on extensional semantics.

table, $|$'empty'$| \cap |$'bottle on the table'$|$, has this property. We need only examine the portion of reality comprising bottles on the table to establish whether the sentence is true.[32]

16.4.2 Absolutist Optimism Revisited

The extensional semantics makes free use of set theory. Barwise and Cooper presuppose the availability of a set-universe M. In the context of standard set theory, this assumption guarantees the existence of all of the sets required as the extensions for predicates, quantifiers and determiners by the extensional semantics.

This assumption poses no problem for the expansionist. In her view, interpretations of languages, including natural languages, form an open-ended hierarchy. Given a version of English, or a formal approximation of the same, as the object language, one can always come to speak a more liberal language in which (i) the members of the object language's universe are the elements of a set M and (ii) the new universe M is closed under set-theoretic operations, rendering the axioms of ZFCU true under the new interpretation. Taking such a language as the metalanguage provides a congenial setting in which to give an extensional semantic theory for the original language. The expansionist can directly follow Barwise and Cooper in casting the extensional semantic theory in set theory, taking the extensions of predicates to be sets of members of M and the extensions of quantifiers to be sets of sets of members of M; and so on.

Expansionist optimism By her lights, the expansionist can give an extensional semantics for determiners such as EVERY by quantifying in the metalanguage over a universe that expands the object language's universe with the requisite set-extensions.

The absolutist cannot always follow suit. Following Barwise and Cooper in casting the extensional semantics in standard first-order set theory is only an option when the members of the object language's universe form a set. But, to repeat the familiar point, in the crucial case when the object language's universe comprises *absolutely* everything, there is no such set. Without a set-universe M for the object language, the absolutist is also without most of the subsets of M that Barwise and Cooper take to be predicate-extensions and most of the sets of subsets of M that they take to be quantifier-extensions. This leads to a certain amount of semantic pessimism.

Absolutist pessimism By his lights, the absolutist cannot give an extensional semantics for determiners such as EVERY, by quantifying in the metalanguage over

[32] Compare Williamson [48, p. 449]. See Peters and Westerståhl [33, sec. 4.5] for an overview of conservativity.

a universe that expands the object language's universe with the requisite set-extensions in the case when the object language's universe is inexpandable.

The restrictionist has a different problem. She cannot take Barwise and Cooper's semantics as it stands. For quantifiers such as EVERY(THING) have the entire universe M as their domain (this is the only predicate-extension that this quantifier 'lives on'). Instead the extensional semantics must be modified to ensure that quantifiers are always restricted to proper subdomains of the universe. One option, following Glanzberg once more, is to introduce background domain relativity into determiner-extensions. On this account, each context c determines a background domain D^c, which has as subdomains all the domains of quantifiers and the extensions of predicates available in that context. For instance, in place of the earlier extensions of THING and EVERY, we instead have that, in each context c, the extension of THING is the background domain D^c, and for each predicate-extension $\boldsymbol{\eta}$ available within D^c:

$$|\text{EVERY}|_c(\boldsymbol{\eta}) = \{\boldsymbol{\theta} \subseteq D^c \mid \boldsymbol{\eta} \subseteq \boldsymbol{\theta}\}$$

(Compare [15, pp. 50–4].)

With these amendments in place, we avoid unrestricted quantification; in each context c, EVERY(THING) now lives on the background domain D^c. But when it comes to stating the semantics for the whole language, the restrictionist faces much the same problem as she did in the case of Tarski-Davidson style semantic theories. To state the compositionality clauses for predicates and quantifiers in every context, we need to deploy something like the following:

(C-θ) For every context c, for any unary-predicate-extension $\boldsymbol{\theta} \subseteq D^c$ and singular-term-extension $\boldsymbol{\tau} \in D^c$ and any unary predicate θ and singular term τ with $|\theta|_c = \boldsymbol{\theta}$ and $|\tau|_c = \boldsymbol{\tau}$: $|\theta(\tau)|_c = \text{T}$ if and only if $\boldsymbol{\tau} \in \boldsymbol{\theta}$.

(C-q) For every context c, for any quantifier-extension $\boldsymbol{q} \subseteq P(D^c)$ and unary-predicate-extension $\boldsymbol{\theta} \subseteq D^c$ and any quantifier q and unary predicate θ with $|q|_c = \boldsymbol{q}$ and $|\theta|_c = \boldsymbol{\theta}$: $|q(\theta)|_c = \text{T}$ if and only if $\boldsymbol{\theta} \in \boldsymbol{q}$.

But, as in the case of satisfaction clauses for contextually restricted quantifiers, these compositionality clauses succeed in specifying the intended extensions of complex expressions in every context only if the quantifiers used to state them in the metalanguage range over a domain $D^\star$ that contains all of the singular-term-, predicate- and quantifier-extensions available in every background domain D^c. But this cannot be so. For, as before, we can come to a strictly wider background domain $D^{\star\star}$, containing singular-term-, predicate- and quantifier-extensions not contained in $D^\star$.

Restrictionist pessimism By her lights, the restrictionist cannot give an extensional semantics for context-sensitive determiners such as EVERY by quantifying in the metalanguage over a proper subdomain of the object language's universe.

16.5 Artificial vs. Natural Metalanguages

16.5.1 Superplural Metalanguages

The absolutist has a natural response to the pessimistic conclusion concerning Mostowski-Barwise-Cooper style semantic theories reached in the last section. What the expansionist achieves by drawing on an ontology surpassing that of the object language, the absolutist can achieve with extended ideological resources. In place of a hierarchy of sets built on top of the set-universe of the object language, the absolutist can give an extensional semantics for an object language with a universe he claims to comprise *absolutely* everything in a metalanguage which is also interpreted over this universe, but avails itself of enough additional levels of plural or higher-order quantification.

Here, we shall consider a plural approach, drawing on the plural semantics developed by Rayo [37], but it is readily adapted in the obvious way to languages with higher-order quantification into predicate position.[33] The extensions of singular terms continue to be members of the universe, and the extensions of sentences continue to be truth-values. Rather than encode the extension of a predicate as a set of things that satisfy the predicate, however, Rayo suggests that the contribution it makes to the semantics of the complex expressions in which it occurs may be encoded, plurally, by those zero or more things that satisfy the predicate. Extension terms such as |BOTTLE| and |THING| become plural terms. For example, under the absolutist's intended interpretation, |BOTTLE|—which, to avoid sounding ungrammatical, we might pronounce 'the satisfiers of BOTTLE'—are the things that comprise every bottle and nothing else; |THING|—the satisfiers of THING—are the things that comprise everything [37, sec 9.2.1].

As in the case of 'domain' and 'universe', the absolutist may continue to indulge in singular 'extension'-talk, as a façon de parler. Following Rayo once more, he may gloss his view as follows: each predicate-extension is now encoded as a plurality rather than a set; for instance, the extension of BOTTLE is the plurality of all bottles; the extension of THING is the plurality of all things. But as with 'domain'-talk, such 'extension'- and 'plurality'-talk cannot be taken at face value, as talking about sets or set-like collections that encode the semantic values of predicates. For in crucial cases, such as in the case of THING, the absolutist contends that there is no such collection [37, p. 225]. Rather to engage in such loose 'plurality'-talk is to suggestively misspeak in a way that—the absolutist may hope—serves to pragmatically communicate what can only be literally said in plural terms.

This leaves the extensions of quantifiers and determiners. Here the absolutist may invoke superplural resources surpassing ordinary plural resources. Given the

[33] Plural or higher-order resources have often been called on in order to formulate absolutist-friendly formulations of model theory. See, for instance, McGee [27, 29], Rayo [38], Williamson [48]. Linnebo and Rayo [26] extend such accounts into the transfinite.

apparent lack of such resources in English, the absolutist may gloss such resources by indulging in more loose-talk: a superplurality is to a plurality, much as a plurality is to an object. A plurality is analogous to a class of objects; likewise, a superplurality is analogous to a 2-class, a class of classes of objects. But a plurality is not a special kind of object; likewise, a superplurality is not a special kind of plurality (or a special kind of object). (Compare Rayo [37, p. 227].)

Having replaced the expansionist's use of sets of members of the universe to encode predicate-extensions with—to indulge in 'plurality'-talk—pluralities of members of the universe, the natural parallel move is to replace the expansionist's use of sets of sets of members of the universe to encode quantifier-extensions with superpluralities of pluralities of members of the universe. (Compare Rayo [36, sec. 8]; [37, sec. 9.2.4].) Thus, for instance, the intended extension of SOME(THING) is the superplurality of all pluralities that have at least one member; and the intended extension of EVERY(THING) is the superplurality of all pluralities that have everything in the object language's universe as a member. With a little bit more coding, the absolutist can also encode the extensions of determiners—which the expansionist treats as set-functions mapping set-predicate-extensions to set-quantifier-extensions—as superplurality-functions mapping plurality-predicate-extensions to superplurality-quantifier-extensions.[34] The compositionality clauses for predicates, quantifiers and determiners, may then be straightforwardly reformulated in the superplural metalanguage, by taking $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ to range over plurality-predicate-extensions, $\boldsymbol{q}$ to range over superplurality-quantifier-extensions, and $\boldsymbol{d}$ to range over superplurality-determiner-extensions.

The absolutist is able to achieve what the expansionist achieves with ontology outside the object language's universe with further ideology over it. Our earlier pessimistic thesis is opposed by the following optimistic one.

Absolutist optimism By his lights, the absolutist can give an extensional semantics for determiners such as EVERY, by quantifying over the object language's universe in a metalanguage that enriches the object language's ideology with the requisite superplural resources.

[34] Note first that we may encode a pair of pluralities $\langle xx, yy\rangle$ as, for instance, the plurality comprising the pair $\langle 1, x\rangle$ for each member x of xx and the pair $\langle 2, y\rangle$ for each member y of yy and nothing else. (Compare Linnebo and Rayo [26, app. B.2].) A determiner-extension may then be encoded as a superplurality zzz of such pairs. Each plurality-predicate-extension xx occurring as the left co-ordinate of a plurality-pair in zzz is mapped to the superplurality-quantifier-extension yyy, comprising those pluralities yy such that $\langle xx, yy\rangle$ is a member of the superplurality-determiner-extension zzz.

16.5.2 Semantic Theorising in Natural Language

While the absolutist is certainly entitled to this much optimism, it is important to realise that it serves only to bound rather than to remove our earlier pessimism. Although the absolutist may attempt to gloss his semantics with loose 'plurality'-talk in natural language, when he comes to *state* the semantics he must depart from natural language and move to an artificial metalanguage enriched with superplural resources. The fact remains that the absolutist cannot follow Barwise and Cooper's actual approach of giving the extensional semantics for an object language he claims to have an *absolutely* all-inclusive universe in a version of English quantifying over the requisite set-extensions.

This comes at the cost of committing the absolutist to an error-theoretic stance towards extensional semantics as carried out by extensional semanticists. Barwise and Cooper are just two examples of semanticists who attempt to uncover the semantic properties of English determiners, working with extensional semantics against a background of first-order set theory.[35] Granted that versions of English have inexpandable universes, as the absolutist claims, this approach misfires. Consider, again, the hypothesis that natural language determiners are conservative: so that the quantifier $d(\eta)$ has an extension that lives on the set $\boldsymbol{\eta}$, the extension of the nominal predicate, in the sense outlined above. Linguists regard this hypothesis as well-confirmed. Yet, as stated by Barwise and Cooper, quantifying over set-extensions, the absolutist must reject it. In his view, in unrestricted contexts, quantifiers like EVERY(THING) live on no set-sized domain.

The absolutist may be inclined to shrug off such an error-theory as not especially costly. He may contend that formulating the semantics in set theory is a mere mathematical convenience that has no bearing on the semantic content of the theory. The philosophically misguided trappings of set-extensions can be skimmed off without loosing the core linguistic insights of the theory. The superplural semantic theory, he may add, does just this.

There is, however, an important limitation to the superplural approach. If, as would seem to be the case, versions of English lack the superplural resources required to give the superplural formulation of extensional semantics outlined in the last section, then, by the absolutist's lights, we cannot give an extensional semantics for significant quantificational fragments of English interpreted over the inexpandable universe using English as the metatheory. Before we can come to semantically reflect in the extensional style on quantificational fragments of our present language we must first come to speak a language with significantly greater ideological resources. Lacking the ability to quantify superplurally in languages they already speak, it would seem that the only way for semanticists to state such a semantic theory is to first learn superplural quantification using the same direct

[35] See also, for instance, Keenan [17], Keenan and Stavi [18], Westerståhl [47], and Peters and Westerståhl [33].

method that they used to master singular and plural quantification, and then to state the theory pragmatically gestured towards in the last section. And this leads to a conclusion about the limits of semantic theorising which does seem to merit the label of pessimism: even if we are prepared to purge extensional semantic theories of set-extensions, semanticists cannot give extensional semantic theories for quantificational fragments of English in English.

Absolutist pessimism By his lights, the absolutist *cannot* give Mostowski-Barwise-Cooper-style semantic theories for significant quantificational fragments of English using English as the metalanguage.

There are two natural strategies the absolutist may employ to contest this conclusion. The first is to seek a less ideologically profligate encoding of extensions so as to avoid the need for superplural resources in the metalanguage. We have already seen that, by the absolutist's lights, sets are ill-suited to this purpose. Might we instead encode each predicate-extension $\boldsymbol{\theta}$ as an object $\hat{\boldsymbol{\theta}}$, but not necessarily a set? As we have already seen in the case of model theory, however, such an approach is immediately scotched by a plural version of Cantor's theorem. The trouble—to indulge in more loose talk—is that there are more plurality-predicate-extensions than there are objects. As Rayo observes, there is no mapping $\hat{\cdot}$ from pluralities to objects such that (i) each plurality-predicate-extension $\boldsymbol{\theta}$ is assigned an object-extension $\hat{\boldsymbol{\theta}}$ and (ii) any two distinct plurality-predicate-extensions $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are assigned distinct object-extensions $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ [37, pp. 224–5]. Parallel reasons force us to ascend to the superplural to capture arbitrary quantifier-extensions. There is no mapping $\tilde{\cdot}$ from superpluralities to pluralities such that (i) each superplurality-quantifier-extension $\boldsymbol{q}$ is assigned a plurality-quantifier-extension $\tilde{\boldsymbol{q}}$ and (ii) any two distinct superplurality-quantifier-extensions $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are assigned distinct plurality-quantifier-extensions $\tilde{\boldsymbol{q}}_1$ and $\tilde{\boldsymbol{q}}_2$. We have resorted to loose-talk to attempt to convey these claims. Indeed, in the second case, given that English lacks the requisite superplural resources, this is our only option in natural language. But each of these claims can be rigorously stated and proved in a suitable superplural setting.[36]

In each case, condition (ii) is non-negotiable. A semantic theory that assigns non-coextensive expressions the same extension is a false theory. But might the absolutist seek to relax condition (i)? Rather than encode arbitrary plurality-predicate-extensions and arbitrary superplurality-quantifier-extensions as objects, the absolutist could restrict himself to just encoding some extensions, for instance, just encoding extensions actually instantiated by the object language under study. Of course, the details of such an encoding remain to be given, and it is not obvious that they will be straightforward: sets remain ill-suited to encode extensions since many *instantiated* plurality-extensions, such as the extension of THING, are not set-sized. But setting the technical details aside, such a response seems to embrace semantic pessimism rather than avoid it. Consider again the compositionality clause

[36] See Rayo [36, sec. 4] and Linnebo and Rayo [26, app. B] for further details. Compare Shapiro [42, thm 5.3].

for quantifier-extensions. To restrict this clause just to a minority of predicate-extensions $\boldsymbol{\theta}$ fails to do justice to the full extent of compositionality present in natural language. The present meaning of the quantifier ‘everything’ determines the truth-conditions for sentences of the form ‘everything θs’ not just for those sentences in the present version of the language, but also for those sentences in any possible extension of the language with new predicates. Upon learning a new predicate, we do not need to check that its extension composes with quantifier-extensions in the standard way. Our commitment to compositionality extends beyond the limits of our present lexicon. To capture this in full generality, therefore, we need to quantify over every predicate-extension (as we did in Sect. 16.4.1) not just a select few.

The second strategy the absolutist may employ to counter the pessimistic conclusion that semantics for English cannot be carried out in English goes in the opposite direction. Rather than attempting to show that superplural resources unavailable in English are not required to give an extensional semantics, the absolutist may instead argue that superplural resources required to give such a semantics are available in English.

There are some putative examples of superplural terms in English. Øystein Linnebo and David Nicolas [25, p. 193] give the following example.

(3) The square things, the blue things and the wooden things overlap.

The sentence has a natural plural reading in which ‘the square things, the blue things and the wooden things’ is a plural term denoting the things that are either square, blue or wooden, so (3) says (truly or falsely) of these things that they overlap. Linnebo and Nicolas argue that a superplural reading is also available, so that (3) says what we might attempt to communicate by saying that the superplurality comprising exactly the plurality of square things, the plurality of blue things and the plurality of wooden things is such that its constituent pluralities overlap in the sense that some object is a member of each of its constituent pluralities.

For the sake of argument, let us grant that such superplural terms are available in English. Their availability falls short of the superplural resources beyond the plural required to give an extensional semantics for English quantifiers in English in two respects. First, while English abounds with both finite plural terms that denote finitely many things such as ‘the members of the Labour party’ and infinite plural terms that denote infinitely many things such as ‘the natural numbers’ or ‘the points in spacetime’, examples in the style of (3) concatenate finitely many plural terms to form a putative superplural term, which is finite in an analogous manner. But to recast an extensional semantics in which quantifier-extensions are often infinite sets of sets requires both terms that denote infinite superpluralities and quantification over the same. No evidence of such terms or quantification in English has yet been forthcoming.

Second, the availability of superplural terms and quantifiers in English would present a theoretical burden as well as providing a metatheoretical resource. The extensional semantics outlined in Sect. 16.4.1 was for a formal language approximating a *singular* fragment of quantificational English. On the parallel treatment for plural expressions, the expansionist may take the extension of a plural term to be a set of elements of the universe M, the extension of a plural predicate to be a set of plural-term-extensions (a set of sets of members of M), and the extension of a plural quantifier to be a set of plural-predicate-extensions (a set of sets of sets of members of M). The absolutist may once again do away with set-extensions by appealing to plural resources. On a plural approach—to speak loosely—the extension of a plural term is a plurality of members of M; the extension of a plural predicate is a superplurality of plurality-plural-term-extensions; and the extension of a plural quantifier is a supersuperplurality—for short, a super^{2}plurality—of superplurality-plural-predicate-extensions.

Such an approach generalises in the obvious way to super^{n}plural terms, predicates and quantifiers. The extension of a super^{n}plural term is a super^{n}plurality; the extension of a super^{n}plural predicate is a super^{n+1}plurality; and the extension of a super^{n}plural quantifier is a super^{n+2}plurality. The Barwise and Cooper approach to quantifiers, consequently, requires a metalanguage two ranks above the object language. Moreover, the Cantorian considerations outlined above generalise in the natural way to show that no ideologically leaner encoding is available.

The upshot of this is that the availability of sufficient superplural resources in English to encode singular-quantifier-extensions does not suffice to dispel absolutist pessimism about the prospects for giving an extensional semantics for English in English. For unless super^{3}plural resources are also available, we shall be unable to encode the extensions of superplural quantifiers. These in turn call for further resources to encode their extensions. Consequently, the absolutist cannot hope to plurally encode the extensions of all of the quantifiers he claims to be available in English unless English possesses super^{n}plural quantification for every finite n. Examples like (3) give us no reason to think that English possesses infinitely many levels of plural quantification beyond the superplural.

16.6 Conclusion

Williamson's conclusion that considerations concerning semantic pessimism favour absolutism over relativism should be tempered twice over. The interaction between absolute generality and semantic pessimism is sensitive both to the variety of relativism in question and the kind of semantic theory under consideration.

References

1. Bach, K. 2000. Quantification, qualification and context: A reply to Stanley and Szabo. *Mind & Language* 15(2–3): 262–283.
2. Barwise, J., and R. Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2): 159–219.

3. Boolos, G. 1984. To be is to be a value of a variable (or to be some values of some variables). *The Journal of Philosophy* 81(8): 430–449, reprinted in Boolos [6].
4. Boolos, G. 1985. Nominalist platonism. *The Philosophical Review* 94(3): 327–344, reprinted in Boolos [6].
5. Boolos, G. 1993. Whence the contradiction? *Aristotelian Society Supplementary Volume* 67: 213–233, reprinted in Boolos [6].
6. Boolos, G. 1998. *Logic, logic, and logic*. Cambridge: Harvard Univesity Press.
7. Cartwright, R. 1994. Speaking of everything. *Noûs* 28(1): 1–20.
8. Davidson, D. 1967. Truth and meaning. *Synthese* 17(1): 304–323.
9. Dummett, M. 1983. *Frege: Philosophy of language*, 2nd ed. London: Duckworth.
10. Dummett, M. 1991. *Frege: Philosophy of mathematics*. London: Duckworth.
11. Evans, G., and J. Altham. 1973. The causal theory of names. *Proceedings of the Aristotelian Society*, Supplementary Volumes 47: 187–225.
12. Fine, K. 2006. Relatively unrestricted quantification. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 20–44. Oxford: Clarendon Press.
13. Frege, G. 1891. *Funktion und Begriff*. English edition: 1980. Function and Concept. Translations from the Philosophical Writings of Gottlob Frege, 3rd ed, ed. P. Geach and M. Black. Oxford: Blackwell.
14. Glanzberg, M. 2004. Quantification and realism. *Philosophy and Phenomenological Research* 69(3): 541–572.
15. Glanzberg, M. 2006. Context and unrestricted quantification. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 45–74. Oxford: Clarendon Press.
16. Hellman, G. 2006. Against 'absolutely everything'! In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 75–97. Oxford: Clarendon Press.
17. Keenan, E. 1984. A Boolean approach to semantics. In *Truth, interpretation and information*, ed. J.A.G. Groenendijk, T.M.V. Janssen, and M.J.B. Stockhof, 65–98. Dordrecht: Foris.
18. Keenan, E., and J. Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9(3): 253–326.
19. Lavine, S. 2006. Something about everything: Universal quantification in the universal sense of universal quantification. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 98–148. Oxford: Clarendon Press.
20. Lewis, D.K. 1970. General semantics. *Synthese* 22(1): 18–67.
21. Lewis, D.K. 1991. *Parts of classes*. Oxford: Basil Blackwell.
22. Lindström, P. 1966. First order predicate logic with generalized quantifiers. *Theoria* 32: 186–195.
23. Linnebo, Ø. 2006. Sets, properties, and unrestricted quantification. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 149–179. Oxford: Clarendon Press.
24. Linnebo, Ø. 2013. Plural quantification. In *The Stanford encyclopedia of philosophy*, Spring 2013 ed, E. Zalta. Stanford: Center for the Study of Language and Information.
25. Linnebo, Ø., and D. Nicolas. 2008. Superplurals in English. *Analysis* 68(3): 186–197.
26. Linnebo, Ø., and A. Rayo. 2012. Hierarchies ontological and ideological. *Mind* 121(482): 269–308.
27. McGee, V. 1997. How we learn mathematical language. *The Philosophical Review* 106(1): 35–68.
28. McGee, V. 2000. Everything. In *Between logic and intuition: Essays in honor of Charles Parsons*, ed. G. Sher and R. Tieszen. Cambridge: Cambridge University Press.
29. McGee, V. 2006. There's a rule for everything. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 179–202. Oxford: Clarendon Press.
30. Mostowski, A. 1957. On a generalization of quantifiers. *Fundamenta Mathematicae* 44(1): 12–36.
31. Parsons, C. 1977. What is the iterative conception of set? In *Proceedings of the 5th international congress of logic, methodology and philosophy of science 1975, Part I: Logic, foundations of mathematics, and computability theory*, ed. R. Butts and J. Hintikka. Dordrecht: Reidel.

32. Parsons, C. 2006. The problem of absolute universality. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 203–219. Oxford: Clarendon Press.
33. Peters, S., and D. Westerståhl. 2006. *Quantifiers in language and logic*. Oxford: Clarendon Press.
34. Quine, W.V.O. 1973. *The roots of reference*. La Salle: Open Court.
35. Quine, W.V.O. 1974. *Methods of logic*. London: Routledge and Keegan Paul.
36. Rayo, A. 2002. Word and objects. *Noûs* 36(3): 436–464.
37. Rayo, A. 2006. Beyond plurals. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 220–254. Oxford: Clarendon Press.
38. Rayo, A., and G. Uzquiano. 1999. Toward a theory of second-order consequence. *Notre Dame Journal of Formal Logic* 40(3): 315–325.
39. Rayo, A., and G. Uzquiano (ed.) 2006. *Absolute generality*. Oxford: Oxford University Press.
40. Russell, B. 1906. The theory of implication. *American Journal of Mathematics* 28(2): 159–202.
41. Russell, B., and A. Whitehead. 1925. *Principia mathematica*, vol. 1, 2nd ed. Cambridge: Cambridge University Press.
42. Shapiro, S. 1991. *Foundations without foundationalism: A case for second-order logic*. Oxford: Oxford University Press.
43. Shapiro, S., and C. Wright. 2006. All things indefinitely extensible. In *Absolute generality*, ed. A. Rayo and G. Uzquiano, 255–304. Oxford: Clarendon Press.
44. Stanley, J., and Z. Szabó. 2000. On quantifier domain restriction. *Mind & Language* 15(2–3): 219–261.
45. Stanley, J., and T. Williamson. 1995. Quantifiers and context-dependence. *Analysis* 55(4): 291–295.
46. Tarski, A. 1935. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica* 1: 261–405, English edition: 1983. The concept of truth in formalized languages. In *Logic, semantics and metamathematics*, 2nd ed, ed. J. Corcoran. Indianapolis: Hackett.
47. Westerståhl, D. 1985. Logical constants in quantifier languages. *Linguistics and Philosophy* 8(4): 387–413.
48. Williamson, T. 2003. Everything. *Philosophical Perspectives* 17(1): 415–465.

Chapter 17
Necessarily Maybe: Quantifiers, Modality and Vagueness

Alessandro Torza

Abstract Languages involving modalities and languages involving vagueness have each been thoroughly studied. On the other hand, virtually nothing has been said about the interaction of modality and vagueness. This paper aims to start filling that gap. Section 17.1 is a discussion of various possible sources of vague modality. Section 17.2 puts forward a model theory for a quantified language with operators for modality and vagueness. The model theory is followed by a discussion of the resulting logic. In Sect. 17.3, the framework will permit us to address a puzzle raised by Elizabeth Barnes and Robert Williams.

The philosophical literature abounds with works on the semantics and logic of modality, and the same can be said of the semantics and logic of vagueness. It comes as a surprise, therefore, that virtually no study is available concerning the interaction of modality and vagueness—especially since the interaction of multiple kinds of modality have been studied quite extensively.[1]

The goal of the present paper is to start filling that gap. Section 17.1 is a discussion of vague modal statements, with a specific focus on the different sources of indeterminacy. By far the most interesting and least dealt with case, as it turns out, is whether a modal statement could be vague as a result of modality's being itself vague. It will be argued that it can, and that an implicit and unexpected defense of such a thesis is to be found in David Lewis' modal realism. Section 17.2 puts forward a model theory for a first-order language featuring both operators expressing metaphysical modality and operators for semantic vagueness. The

This work has been made possible in part by the CONACyT grant CCB 2011 166502 and the PAPIIT grant IA400412.

[1]For instance, Segerberg [15], Thomason [18], Gabbay [4].

A. Torza (✉)
Instituto de Investigaciones Filosóficas, UNAM, Circuito Mario de la Cueva, Ciudad Universitaria, Del. Coyoacán, México D.F. 04510, Mexico
e-mail: atorza@me.com

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_17

interpretation of metaphysical modalities is based on counterpart theory, whereas semantic vagueness is understood in terms of precisifications. The definition of the model theory is followed by a discussion of the resulting logic. In Sect. 17.3, the framework will permit us to settle an open question. Barnes and Williams [1] have claimed that a language combining expressions for both vagueness (modeled via precisifications) and modality (modeled via possible worlds) would obey an overly revisionary logic, namely by making inconsistencies satisfiable. I will argue that the claim is unwarranted.

17.1 Modal Vagueness

This section is a critical examination of the ways in which modal notions could be vague. By ‘modality’ I here mean *metaphysical* modality, unless otherwise stated. In particular, I assume that metaphysical modalities are *absolute*, in the sense that, if it is possible that p, in any sense of ‘possible’, then it is metaphysically possible that p. For instance, since quantum teleportation is physically possible, then it is also metaphysically possible. Likewise, since it is not a historical necessity that the Archduke Franz Ferdinand of Austria had to be killed in Sarajevo, then it wasn’t metaphysically necessary, either. On the other hand, the physical impossibility of superluminal causation need not be understood as a metaphysical impossibility. When possibility is construed in terms of existential quantification over worlds, absolute possibility is *unrestricted* existential quantification over worlds. Relative possibility is restricted existential quantification over worlds.[2]

17.1.1 Predicates

First of all, a modal statement can be vague by containing a vague *predicate*, simple or complex. Given a modal language, I take the semantic value of a predicate to be a set of *possibilia*, and the semantic value of a predicate at world w to be the restriction to w of its semantic value. Now, consider a community of sloppy chemists whose use of the term ‘hydrogen’ is indeterminate between two precise meanings: the element with atomic number 1 vs. an isotope of the element with atomic number 1 which has actually been observed. Since no isotope of hydrogen has ever been observed (in nature or in a lab) with more than six neutrons (viz., hydrogen-7), the following statement is semantically vague in sloppy-chemistese:

1. No hydrogen atom could possibly have seven neutrons

[2] On absolute modalities, see Hale [6].

For, there is one sense of 'hydrogen' in sloppy-chemistese—the one agreeing with our own use of the term—which allows hydrogen atoms to have more neutrons than have ever been observed, and another sense which excludes such a possibility.

It is noteworthy that the occurrence of a vague predicate, simple or complex, in a modal statement will not automatically make that statement vague—just like, in general, the occurrence of a vague expression in a statement need not make the latter vague. To wit, it can be vague whether

2. Zach is bald

and yet definitely true that

3. It is contingent whether Zach is bald

In order to see that, suppose there is a range of precisifications 'bald_1', … 'bald_m', 'bald_{m+1}', … 'bald_n', such that Zach is 'bald_m' but not 'bald_{m+1}'. In this scenario, (2) will indeed be vague. But, as long as logical space is sufficiently plentiful, for every $i \leqslant n$,

4. It is contingent whether Zach is bald_i

Hence, (3) is definitely true.

17.1.2 Intensional Identity

A further scenario is one in which modal statements are vague due to the nature of *intensional identity*. I will draw on Lewis [9] in construing intensional identity in terms of a counterpart relation, in such a way that 'x is possibly P' is paraphrased as 'for some world w, the counterpart of x at w is P'. (I make no mention of accessibility here, since modality is taken to be absolute.) The counterpart of x at w is the individual which best represents x at w in terms of content and context.[3]

Here is an example of vague intensional identity. Consider a world of one-way eternal recurrence w such that each epoch is a duplicate of the history of the actual world. Insofar as w contains duplicates of actual Socrates (in fact, one for each epoch), the possibility of such a world makes it intuitively true that

5. Socrates could have lived in a world of one-way eternal recurrence so-and-so

where 'so-and-so' is short for the above description of w. But if the actual world had been w, in which epoch would have Socrates lived? It seems sensible to say that in some sense he could have lived in the first epoch, in some sense he could have lived in the second, etc. One way to accommodate this intuition within counterpart theory is to admit the existence of infinitely many duplicate worlds w_1, w_2,… of w,

[3] We can safely assume that the counterpart relation is reflexive. Unlike Lewis, I assume throughout that nothing has multiple counterparts at a world. I expect my choice to make sense in light of the following remarks.

such that in w_1 the counterpart of actual Socrates is the Socrates-duplicate in the first epoch, in w_2 the counterpart of actual Socrates is the Socrates-duplicate in the second epoch, etc. Each of the following will then have to be true:

6.1. Socrates could have lived in the first epoch of a world of one-way eternal recurrence so-and-so
6.2. Socrates could have lived in the second epoch of a world of one-way eternal recurrence so-and-so
etc.

Needless to say, each w_i will make (5) true as well. Nevertheless, this solution to the above *desideratum* entails haecceitism, since there will be worlds (infinitely many, in fact) that differ in a merely non-qualitative way, viz., with respect to which of the Socrates duplicates happens to be Socrates. Since not everybody is a friend of haecceitism, it would be desirable to accommodate the above intuition in a way that does not entail such a metaphysical position. Here is how. When we say that in some sense Socrates could have lived in the nth epoch of a world like w, for every n, in counterparts theory we do not have to express such scenarios by means of possibilities. We could instead mean something different, namely that for every Socrates-duplicate s_n in w, there is a way of making the counterpart relation precise that picks s_n out. Hence, with respect to the one and only w, one precisification of the counterpart relation associates actual Socrates to the Socrates-duplicate s_1 in the first epoch, another precisification of the counterpart relation associates actual Socrates to the Socrates-duplicate s_2 in the second epoch, etc. Each of the following will then be *vague*:

6.1. Socrates could have lived in the first epoch of a world of one-way eternal recurrence so-and-so
6.2. Socrates could have lived in the second epoch of a world of one-way eternal recurrence so-and-so
...

On the other hand, (5) will be true as per the original intuition, since it remains true under every precisification of the counterpart relation.[4]

17.1.3 Quantifiers

A third potential source of modal vagueness are *quantifier-like expressions*. Garden-variety modal languages feature two kinds of quantifier-like expressions: modal

[4]The problem of vague intensional identity is reminiscent of the well-known problem of relative intensional identity discussed in Lewis [10], [11, p. 248], Gibbard [5], Stalnaker [17]. The crucial difference between the two cases is that in the latter, but not in the former, fixing the context of utterance suffices to specify a counterpart relation. Many thanks to Maite Ezcurdia for helping me see this distinction.

operators, ranging over worlds, and first-order variable-binding quantifiers, ranging over world-bound individuals. I will now argue that there are indeed cases in which the vagueness of modal statements stems from indeterminacy about what worlds or individuals there are.

Before proceeding, it is important to clarify one issue. First of all, quantification over worlds or world-bound individuals in the background language of counterpart theory can be restricted or absolute. In the case of world-bound individuals, a restricted quantifier is defined by an unrestricted quantifier and a sortal predicate. We will deal with vague unrestricted quantifiers in due course, whereas sortal predicates can be broken down to simpler constituents. Therefore, the case of vague restricted quantifiers does not need to be treated separately. As to restricted quantification over worlds, that expresses relative modality. Since we are only concerned with absolute (metaphysical) modality, this case is irrelevant for present purpose.

On the present counterpart-theoretic approach, what there is falls into two categories: worlds and world-bound individuals. Let's first consider the case in which a modal statement is vague because the domain of world-bound individuals is vague. Call *dyadism* the thesis that there are exactly two objects.[5] It should not be too controversial that dyadism is false. But is it at least possible? In other words, I am considering whether the following modal statement is true:

7. There could have been exactly two objects[6]

The answer will depend, among other things, on the underlying mereology. On the one hand, the mereological universalist believes in unrestricted composition. The range of her quantifier will therefore be closed under arbitrary fusions. In this sense of 'there is', it is impossible for there to be exactly two objects, provided that worlds are closed under fusions. At the other end of the mereological spectrum is the nihilist, denying the existence of proper parts and for whom a quantifier can only range over mereological atoms.[7] On the latter sense of 'there is', dyadism is possible in virtue of the existence of a world containing exactly two mereological atoms. So, as long as it can be indeterminate which mereology constrains our quantifiers, (7) will be vague. Notice that in the present case it is not vague what worlds there are, and yet it is vague what individuals there are at each world.[8] The moral is that modal vagueness can ensue if quantification over world-bound individuals is unsharp.

---

[5]Dyadism is modeled after monism, the thesis that there is exactly *one* object. Monism, which has famously been defended by Parmenides, should not be conflated with *priority monism*, the view that the world is prior to its parts, as advocated recently in Schaffer [14].

[6]This sentence can be regimented in purely-logical first-order vocabulary: $\Diamond\exists x\exists y(\sim x = y \wedge \forall z(z = x \vee z = y))$.

[7]For the sake of simplification, I am ignoring here the possibility of gunk.

[8]The whole discussion should be rephrased in terms of *concrete* individuals, if worlds are assumed to be closed under set-theoretic constructions or, more generally, if an infinite number of abstracta exists by necessity. Otherwise, both the universalist and the mereological nihilist will regard (7) as trivially true.

17.1.4 Modals

I now turn to the case in which operators expressing absolute modality are vague, which is the result of its being indeterminate what worlds there are. Consider the following example.[9] Let k be a universal constant occurring in some physical equation $E(k)$ and satisfying the following two conditions:

(i) The value of k is *contingent*, so that there is a range of possible worlds which are obtainable by varying k in E
(ii) The range of k is *bounded*, which is to say, there is an interval of possible values of k

Now, if i is a value of k outside the interval, there exists a series from the actual value to the impossible value i. Consider a scenario in which scientists are unable to identify a sharp cutoff in the series. As a result, there ought to be some j between i and the actual value which is neither definitely possible nor definitely impossible. Consequently, it will be indeterminate whether there are worlds in which k takes on value j. Since this scenario makes it indeterminate what worlds there are, the range of modal operators will be vague. The statement

8. It is possible that both $E(k)$ and $k = j$

will then have to be vague, in virtue of its being true on some but not all senses of 'possible'.

One might object that (8) can be interpreted as vague, but not in the intended way. For our goal was to show how modal vagueness can be traced to metaphysical modality itself. However, goes the objection, metaphysical modality is *absolute*, whereas the above example could equally be interpreted as providing an instance of vague *relative* modality. In order to see that, we could rephrase the story as follows. Let's assume that absolutely every value of k is metaphysically possible, and yet there is an interval of k-values which defines the *physically* possible worlds (i.e., possible relative to the physical equation $E(k)$). Statement (8) would then be definitely true, when 'possible' is unrestricted. But if we regard 'possible' as expressing physical modality, then (8) will be vague.[10] The moral of the objection is that the story is underspecified. Unless we have independent reasons to rule out certain k-values as absolutely, rather than merely physically impossible, the above story is compatible with the weaker claim that physical modality is vague.

Whether the story could be further specified so as to avoid the above charges depends on our criteria for discriminating physical from metaphysical possibility, and in particular for identifying metaphysically possible worlds. Instead of replying directly to the objection, I will consider a new story which also aims to show that

[9]Something in the vicinity of this was suggested to me by Daniel Berntson in private conversation.

[10]If relative modalities are expressed model-theoretically by means of accessibility relations, as is customary, vague physical modality would then be modeled by an appropriately vague accessibility relation defined over a sharp domain of metaphysically possible worlds.

modality is vague, but which does not underdetermine whether the modality at issue is absolute or relative.

In order to guarantee that logical space be sufficiently plentiful, it is routine to assume the so-called *principle of plenitude*:

PL: Absolutely every way the world could be is a way a world is

However, the modal realist cannot appeal to such a principle, and for a simple reason. Since in modal realism ways a world is or could be are identified with worlds, **PL** would be tantamount to the logical truth: absolutely every world is a world.[11] Lewis responded by trying to capture plenitude with a *principle of unrestricted recombination*, which roughly says that every distribution of natural properties in space-time constitutes a world. The principle, on its intended application, entails

UR: For any objects in any worlds, there exists a world that contains any number of duplicates of all those objects[12]

The idea behind **UR** is that logical space should be closed under the operation of patching together copies of arbitrary collections of *possibilia* in a single world.

As it turns out, however, **UR** leads to paradox and therefore the modal realist cannot rely on it as a replacement for **PL**. The first *reductio* of **UR** was offered in Forrest and Armstrong [3], where it is argued that the principle is inconsistent with the assumption that the *possibilia* form a set. Nolan [12] has shown that, although the Forrest-Armstrong argument is invalid, a new and simpler proof is available, which goes as follows. Let k be the cardinality of the set of all *possibilia*. If a is an object, by **UR** there exists some world w containing 2^k duplicates of a. But $k < 2^k$ and yet the objects existing at w are a subset of all the *possibilia*. So, some *possibilia* are more than the whole. Hence, the *reductio*.

There are two main strategies available for blocking Nolan's argument against **UR**. First, we could simply assume that the collection of all *possibilia* forms a proper class. In that case, there would be no cardinal k measuring its size, and the *reductio* would not go through. This is the road taken and defended by Nolan himself.[13] There is nevertheless a number of reasons for resisting the prospects of a class-sized universe. Probably the most obvious reason is that, since the modal realist identifies properties and relations with sets, and since proper classes are not members of any set, then proper class-sized properties will not have any second-order properties or relations. For instance, if the property of having mass is proper class-sized, then we won't be able to say of that property that it is natural. In fact, there would be no (adequate) set-theoretic representative of naturalness and, therefore, no property of naturalness at all! This is of course a very unsavory outcome for the Lewisian. Nolan's solution is to identify properties with universals, and second order properties with sets of universals. Although this approach reinstates the

[11]Lewis [11, p. 86].

[12]Lewis [11, p. 88], Nolan [12, p. 239].

[13]Nolan [12, p. 248].

existence of all second-order properties, as desired, it comes at the cost of depriving modal realism of one of its main theoretical virtues, viz., its capacity to provide a nominalistic theory of properties.[14]

I now turn to the second strategy for resisting Nolan's *reductio*, which is to weaken **UR**. As it turns out, this option will produce the instance of modal vagueness that we are seeking. A way of restricting recombination, which was put forward by Lewis [11, p. 89] is to assume that there exists a least upper bound to the number of objects which could coexist in any single world by virtue of some spacetime constraint. Accordingly, the *principle of restricted recombination* states that every distribution of natural properties in space-time constitutes a world, *shape and size permitting*.[15] On its intended construal, the principle entails:

RR: For any objects in any worlds, there exists a world that contains any number of duplicates of all those objects, shape and size permitting.

But what does 'shape and size permitting' mean? Here is what the modal realist could say:

> My thesis is existential: there is *some* break, and the correct break is sufficiently salient within the mathematical universe not to be *ad hoc*. If study of the mathematical generalisations of ordinary spacetime manifolds revealed one salient break, and one only, I would dare to say that it was the right break—that there were worlds with all the shapes and sizes of spacetime below it, and no worlds with any other shapes and sizes. If study revealed no suitable breaks, I would regard that as serious trouble. If study revealed more than one suitable break, I would be content to profess ignorance—incurable ignorance, most likely. (Lewis [11, p. 103])

There are three scenarios opening up. The first case is that mathematics tells us that there is exactly one way of generalizing ordinary spacetime manifolds—for example by admitting all and only manifolds of finitely many dimensions, or perhaps up to countably many dimensions (in which case, a world could contain at most continuum many wholly distinct objects). The second case is that ordinary spacetime manifolds can be generalized to manifolds with any arbitrary number of dimensions. This is the problematic scenario mentioned by Lewis, where **RR** would collapse into **UR**. The third case is that there are multiple and equally natural ways of generalizing ordinary manifolds, but inducing different least upper bounds to the number of possibly coexistent objects.

We are interested in the third scenario: the least upper bound exists, but there are multiple equally natural candidates for playing that role. In that case it would be indeterminate which *particular* value is the least upper bound. (The failure of existential instantiation is a red flag that we are here dealing with vagueness.) Sup-

[14]For the sake of completeness, I should mention that Nolan in fact proposes a second solution which does not involve commitment to universals. However, this alternative approach requires that "all and only the natural properties possess singletons". It is questionable whether the extent of set theory should be sensitive to such metaphysical distinctions, especially since it is unclear whether there is a sharp cutoff for the (perfectly) natural properties.

[15]Cf. Divers [2, p. 102]

pose for instance that there are two suitable generalizations of ordinary spacetime manifolds, which induce k_1 and k_2 as least upper bounds, respectively. Assuming that $k_1 < k_2$, it will then be vague whether there is a world comprising k_2-many wholly distinct objects. Now, let's make the auxiliary assumption that there is no quantifier vagueness, which is to say, that the number of objects existing at any given world is sharp. It follows that it is vague what worlds there are, absolutely. For at one precisification there will be some world w_{k_2} comprising k_2-many objects; and since the size of worlds is definite, w_{k_2} will not exist at the precisification which allows at most k_1 coexistent objects. We can conclude that modality can be vague in virtue of what worlds there are, absolutely.

A related issue must be raised at this juncture. I just argued that quantifier expressions in the modal language, namely quantifiers proper and modals, can be vague. I have done so by exhibiting cases in which (i) quantification over worlds and world-bound individuals in the language of counterpart theory is vague and (ii) such quantifiers are absolute, as they range unrestrictedly over all worlds and *possibilia*. Moreover, I have been assuming throughout that (iii) vagueness is analyzed via precisifications.

My case, however, runs counter to an argument put forward in Sider [16], which aims to show that

(V) If vagueness is given a precisificational account and existence is expressed by the unrestricted existential quantifier, then existence cannot be vague

If the argument for (V) is sound, the above conditions (i)–(iii) are bound to be jointly inconsistent. Nevertheless, Torza [19] has argued that Sider's argument is compatible with a weak form of vague existence. Let us take a closer look at the dialectics.

Sider's alleged proof has the form of a *reductio ad absurdum*. Suppose that

P. $\ulcorner \exists x \phi \urcorner$ is vague

(where ϕ is precise). As long as $\exists$ is absolute and vagueness is construed via precisifications, it can be shown that (P) entails an inconsistency. At this point Sider applies *reductio* and infers that (P) is false. As remarked in Williamson [20, p. 152], however, *reductio ad absurdum* is valid for bivalent languages. In this particular case, therefore, we may infer the falsity of (P) if the metalanguage of $\ulcorner \exists x \phi \urcorner$ is precise. But notice that (P) is equivalent to

P'. In some precisification $\ulcorner \exists x \phi \urcorner$ is true and in some precisification $\ulcorner \exists x \phi \urcorner$ is false

which involves quantification over precisifications of the language of $\ulcorner \exists x \phi \urcorner$. If the *set of* precisifications is not itself precise, *reductio* may not be applied. All we could infer, then, is that (P) is untrue, i.e., either false or vague—an instance of *weak reductio* (cf. Keefe [8, p. 180]). In order to complete the original *reductio*, Sider would now have to show that (P) is not vague, i.e., that $\ulcorner \exists x \phi \urcorner$ is not second-order vague. Torza [19] shows how to set up a *reductio* of second-order vague existence, Sider style. But if a *reductio* of vague existence presupposes that the metalanguage of the quantifier $\exists$ be precise, likewise a *reductio* of second-order vague existence presupposes that the meta-metalanguage of $\exists$ be precise. And so forth and so on. The upshot is that neither side has the upper hand. In particular, we have no reasons to rule out the possibility that existence be vague at all orders—i.e., vague, and second-order vague, and third order vague, etc. Following Torza [19], I call *super-vague* any instance of quantification which is vague at all orders in this sense. Accordingly, whenever I speak here of vague existence and modality, I actually mean super-vague existence and modality.

Now that the issue of the coherence of vague quantification (albeit in a weak form) has been cleared up, we can conclude that modal languages have at least four possible sources of vagueness: predication, intensional identity, quantifiers and modals. In the next section I turn to the second goal of this paper, namely to work out a model theory for languages containing both modal operators and vagueness operators that accommodates the observation from this section.

17.2 Modal Vagueness, Regimented

17.2.1 Supervaluationary Counterpart Semantics

Modal languages, when sharp, can be interpreted by means of *counterpart models*.[16] If the object language is vague, however, vanilla counterpart models are inadequate. What we need are structures with multiple precisifications, each of which will itself be a counterpart model. While in standard counterpart semantics sentences are evaluated at a world w, in the supervaluationary case we want to evaluate sentences at a pair $\langle s, w\rangle$, where s is a precisification and w a world index. The elements of a model that are allowed to vary across precisifications will define which parts of the modal language are unsharp. Given what has been said in Sect. 17.1, we want non-logical constants to vary across precisifications, so as to allow for vague predication. We want the counterpart relation to vary, too, in order to represent the vagueness of intensional identity. We want the domain of a world to be able to vary across precisifications, if existence is to be vague. Finally, we want the whole set of worlds itself to vary across precisifications, to account for vague modality.

In order to meet the above desiderata, I start by defining a *supervaluationary counterpart frame* (*SC-frame*), which is a structure $\mathcal{F} = \langle Q, @, R, U, Dom, c\rangle$, where

- $Q \subseteq S \times W$, for S, W disjoint sets
- $\langle s, @\rangle \in Q$, for every $s \in S$
- $R \subseteq Q^2$ s.t. $\langle s, w\rangle R\langle s', w'\rangle \rightarrow s = s'$
- U is a set disjoint from S and W
- $Dom : Q \longrightarrow \mathcal{P}(U)$ s.t.

[16] The *loci classici* of semantics based on counterparts are Lewis [9], Hazen [7].

  - If $w \neq w'$, then $Dom(\langle s, w\rangle) \cap Dom(\langle s', w'\rangle) = \emptyset$
  - $U = \bigcup_{\langle s,w\rangle \in Q} Dom(\langle s, w\rangle)$

- $c : U \times Q \longrightarrow U$ s.t.

  - $c(a, \langle s, w\rangle) \in Dom(\langle s, w\rangle)$
  - If $a \in Dom(\langle s, w\rangle)$ and $b = c(a, \langle s', w'\rangle)$, then $s = s'$
  - If $a \in Dom(\langle s, w\rangle)$, then $c(a, \langle s, w\rangle) = a$
  - If $a, b \in Dom(\langle s, w\rangle)$ and $a \neq b$, then $c(a, \langle s, w'\rangle) \neq c(b, \langle s, w'\rangle)$

A few comments are in order. S and W are sets of indices for precisifications and worlds, respectively, in such a way that each coordinate $\langle s, w\rangle$ is identified with a world-in-a-precisification (or, simply, a world). The reason why $\mathcal{F}$ is defined on Q, rather than the whole product-set $S \times W$, is that a world-coordinate w may pick out a world at some precisification s but not at some other s'. This fact captures the idea that the set of worlds, over which unrestricted modal operators range, can be vague.

Each precisification s will feature an actual world $\langle s, @\rangle$.

R is the accessibility relation, which relates worlds to worlds within the same precisification. Since we are interested here in absolute modalities, from now on I will simply assume that R is universal (viz., $\langle s, w\rangle R \langle s, w'\rangle$, for every s, w, w') and omit any reference to it altogether.

U are the individuals.

Dom maps each world $\langle s, @\rangle$ to a set of world-bound individuals, and every individual exists at some world.

The function c assigns to each individual a counterpart at every world within the same precisification, so that distinct world-mates have distinct counterparts at any given world. Notice that the assumption that everything has a counterpart at every world (within the same precisification) is arguably too strong. For instance, it is reasonable to assume that some worlds are so radically different from ours that nothing over there could ever represent, say, actual Socrates. Nevertheless, for the sake of simplicity I will stick to the present choice, with the proviso that, in a fully adequate semantics, an individual may fail to have counterparts at some world.

Now, let $\mathcal{L}$ be a first-order language endowed with identity and an infinite set of n-ary predicate constants, for each $n > 0$. The expansion of $\mathcal{L}$ with the sentential necessity operator $\Box$ (definiteness operator Δ) is referred to as $\mathcal{L}_\Box$ ($\mathcal{L}_\Delta$). The union of $\mathcal{L}_\Box$ and $\mathcal{L}_\Delta$ is $\mathcal{L}_{\Box\Delta}$. In $\mathcal{L}_\Box$ the possibility operator is defined by the condition $\Diamond\phi := {\sim}\Box{\sim}\phi$. In $\mathcal{L}_\Delta$, the 'in some sense' operator ∇ is defined by $\nabla\phi := {\sim}\Delta{\sim}\phi$. The vagueness operator I is defined by $I\phi := \nabla\phi \wedge \nabla{\sim}\phi$.

A *supervaluationary counterpart model* (*SC-model*) is a structure $\mathcal{M} = \langle \mathcal{F}, \sigma\rangle$ where

- $\mathcal{F}$ is a SC-frame
- For every $\langle s, w\rangle \in Q$,

  - $\sigma(=, \langle s, w\rangle)$ is the identity relation over $Dom(\langle s, w\rangle)$
  - $\sigma(P, \langle s, w\rangle) \subseteq Dom(\langle s, w\rangle)^n$, for every n-ary predicate constant P

Given the set *VAR* of variables in a language, a *value assignment for VAR* over $\mathcal{M}$ is a set of partial functions $\{\xi_s\}_{s\in S}$ s.t.

- $\xi_s : VAR \rightarrow Dom(\langle s, @\rangle)$
- $\bigcup_{s\in S} \xi_s$ is a total function $f : VAR \longrightarrow \bigcup_{s\in S} Dom(\langle s, @\rangle)$
- if $\xi_s(x)$ and $\xi_t(x)$ are both defined, then $\xi_s(x) = \xi_t(x)$

The choice of breaking down an assignment for the variables into a set of partial functions aims to capture the idea that, since existence is vague, a variable may or may not successfully refer, depending on a particular precisification.[17]

Local truth, i.e. truth at a world-in-a-precisification $\langle s, w\rangle \in Q$ in $\mathcal{M}$ under $\{\xi_s\}_{s\in S}$ is defined thus:

1. If $\phi = P(x_1 \dots x_n)$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff $c(\xi_s(x_i), \langle s, w\rangle)$ is defined for all $i \in \{1, \dots, n\}$ and $\langle c(\xi_s(x_1), \langle s, w\rangle) \dots c(\xi_s(x_n), \langle s, w\rangle)\rangle \in \sigma(P, \langle s, w\rangle)$
2. If $\phi = \sim\psi$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \nvDash \psi$
3. If $\phi = \psi \wedge \chi$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \psi$ and $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \chi$
4. If $\phi = \forall x\psi$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff, for every $\{\xi'_s\}_{s\in S}$ differing from $\{\xi_s\}_{s\in S}$ at most on x, $(\mathcal{M}, \langle s, w\rangle, \{\xi'_s\}_{s\in S}) \vDash \psi$
5. If $\phi = \Box\psi$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff, for every $\langle s, w'\rangle \in Q$, $(\mathcal{M}, \langle s, w'\rangle, \{\xi_s\}_{s\in S}) \vDash \psi$
6. If $\phi = \Delta\psi$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff, for every $\langle s', w\rangle \in Q$, $(\mathcal{M}, \langle s', w\rangle, \{\xi_s\}_{s\in S}) \vDash \psi$

One issue we were faced with in definition of local semantics is how to evaluate an atomic formula $P(x)$ at a precisification where x is non-referring. The present framework always assigns 'false' to such formulas. As a consequence, local truth defines a *negative free semantics*.[18]

We can finally define the notions of truth-in-a-model, logical consequence and validity as follows.

ϕ is *true in* $\mathcal{M}$ *under* $\{\xi_s\}_{s\in S}$ $((\mathcal{M}, \{\xi_s\}_{s\in S}) \vDash \phi)$ iff, for every $s \in S$, $(\mathcal{M}, \langle s, @\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$.

ϕ is *true in* $\mathcal{M}$ $(\mathcal{M} \vDash \phi)$ iff, for every $\{\xi_s\}_{s\in S}$, $(\mathcal{M}, \{\xi_s\}_{s\in S}) \vDash \phi$.

ϕ is a *consequence* of Γ $(\Gamma \vDash \phi)$ iff, for every $\mathcal{M}$, if $\mathcal{M} \vDash \Gamma$ then $\mathcal{M} \vDash \phi$

ϕ is *valid* $(\vDash \phi)$ iff, for every $\mathcal{M}$, $\mathcal{M} \vDash \phi$

It is noteworthy that SC-frames could be enriched by adding an *admissibility relation* A, where $A \subseteq Q^2$ and $\langle s, w\rangle A \langle s', w'\rangle \rightarrow w = w'$. The truth condition (6) for formulas governed by Δ in a *SC*-model would then have to be revised accordingly:

6′. If $\phi = \Delta\psi$, then $(\mathcal{M}, \langle s, w\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$ iff, for every $\langle s', w'\rangle \in Q$ s.t. $\langle s, w\rangle A \langle s', w'\rangle$, $(\mathcal{M}, \langle s', w'\rangle, \{\xi_s\}_{s\in S}) \vDash \psi$

[17] This definition of an assignment for the variables is developed in Torza [19].

[18] For a motivation and discussion, see Torza [19]. For an elucidation of free logics, see Nolt [13].

In fact, such a revision in the definition of *SC*-models is not only possible but even necessary in the light of what has been said in Sect. 17.1 concerning vague quantification. Indeed, recall that absolute quantifiers can be vague as long as the vagueness extends to all orders, which is to say, as long as the quantifiers are super-vague. Clearly, this idea can be captured only in models that allow for higher-order vagueness. On the other hand, supervaluationary models without an admissibility relation, or in which admissibility is an equivalence relation, do not admit of high-order vagueness, since they validate the schema $I\phi \rightarrow \Delta I\phi$. We must conclude that SC-models in which object-language quantifiers and modals are super-vague require an admissibility relation A which is not reflexive, symmetric and transitive. (Williamson [21] and Torza [19] argue that the most natural approach is to drop transitivity.) Nevertheless, I will refrain from adding the admissibility relation A as suggested, in attempt to simplify the model theory.

17.2.2 Logic

What is the logic of a language $\mathcal{L}_{\Box\Delta}$ whose behavior is defined by SC-semantics? I am going to break down the question into four subproblems. I will first consider a set of $\mathcal{L}$-theses, i.e., schemata and rules of inference which can be formulated in the extensional sub-language $\mathcal{L}$, and check which of them are validated in $\mathcal{L}_{\Box\Delta}$. I will then repeat the test with respect to a set of $\mathcal{L}_{\Box}$-theses, which are the purely modal theses. I will next consider a set of $\mathcal{L}_{\Delta}$-theses, schemas and inference rules that usually hold on a supervaluationary interpretation of $\mathcal{L}_{\Delta}$. Finally, I consider the $\mathcal{L}_{\Box\Delta}$-theses, which can only be formulated in a language with both modal and definiteness operators.

17.2.2.1 $\mathcal{L}$-Logic

Let us establish which schemas and inference rules, which can be formulated in an extensional first-order language $\mathcal{L}$, hold in the expanded language $\mathcal{L}_{\Box\Delta}$.

Let ϕ be a $\mathcal{L}_{\Box\Delta}$-formula. Note that every atomic $\mathcal{L}_{\Box\Delta}$-formula is either locally true or locally false (under an assignment), and that sentential connectives are defined classically. Therefore, if ϕ is a *classical tautology* and $\mathcal{M}$ a SC-model, $(\mathcal{M}, \langle s, @\rangle, \{\xi_s\}_{s\in S}) \vDash \phi$, for every $s \in S$. Hence,

TAUT. $\vDash \phi$, if ϕ is a classical tautology

Moreover, *Modus Ponens* holds:

MP. $\phi, \phi \rightarrow \psi \vDash \psi$

Other classical inference forms, however, are invalid in supervaluationary counterpart semantics. As discussed in Keefe [8], *reductio ad absurdum*, *contraposition*, *conditional proof* and *argument by cases* typically fail in supervaluationism.

Nevertheless, weakened versions of those forms of inference hold in general in supervaluationism and specifically in SC-semantics, namely:

RA. If $\Gamma, \phi \vDash \bot$, then $\Gamma \vDash \sim\Delta\phi$
CON. $\phi \vDash \psi$, then $\sim\psi \vDash \sim\Delta\phi$
CP. If $\Gamma, \phi \vDash \psi$, then $\Gamma \vDash \Delta\phi \rightarrow \psi$
AC. If $\phi \vDash \chi$ and $\psi \vDash \chi$, then $\Delta\phi \vee \Delta\psi \vDash \chi$

A discussion and defense of these quasi-classical inference forms from a supervaluationary point of view is put forward in Keefe [8, p. 179].

Let us now turn to quantified logic. As I had remarked in Sect. 17.1 already, classical *existential instantiation* fails in supervaluationary frameworks. For $\exists x\phi$ can be true at all precisifications, and yet there may be no value of x which makes ϕ true at all of them. This fact remains true in supervaluationary counterpart semantics.

Existential generalization, which is instead a typically valid form of inference in supervaluationary semantics, fails in the present framework, too. For example, for P a non-logical constant, it could be that $\sim P(x)$ is true in a model (under an assignment), whereas $\exists x \sim P(x)$ is untrue. In order to see that, just consider a model in which the variable x is undefined at $\langle s, @\rangle$, for some s, and has a value in the anti-extension of P at $\langle s', @\rangle$, for every other s'. The failure of classical existential generalization is clearly due to the fact that local truth is defined in terms of negative free semantics. As it turns out, it can be proven by induction on the complexity of ϕ that a weaker form of existential generalization, typical of free logics, holds in supervaluationary counterpart semantics:

∃G. $\phi(x), \exists y(x = y) \vDash \exists x\phi(x)$

It is easy to show the *equivalence of self identity and existence*:

EX. $x = x \leftrightarrow \exists y(x = y)$

Note that the first-order axiom $x = x$, expressing the *reflexivity of identity*, fails. However, the weaker, quantified version is valid:

SI. $\forall x(x = x)$

Leibniz' Law is valid in the quantifier-free form:

LL. $x = y \rightarrow (\phi(x) \rightarrow \phi(y))$

In fact, a stronger principle holds:

LL$^+$. $\Diamond\nabla x = y \rightarrow (\phi(x) \rightarrow \phi(y))$

The two laws **LL** and **LL$^+$** can be proved concurrently by induction on the complexity of ϕ.

Proof of **LL** and **LL$^+$**. I will show only the most interesting cases of the induction. Reference to a fixed model $\mathcal{M}$ is left implicit throughout.

1. Let $\phi(x) = P(x)$. Assume $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash P(x)$.

1.1 If $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash x = y$, then $\langle c(\xi_s(x), \langle s, w\rangle), c(\xi_s(y), \langle s, w\rangle)\rangle \in \sigma(=, \langle s, w\rangle)$, i.e., $c(\xi_s(x), \langle s, w\rangle) = c(\xi_s(y), \langle s, w\rangle)$. Since $c(\xi_s(x), \langle s, w\rangle \in \sigma(P, \langle s, w\rangle)$, then $c(\xi_s(y), \langle s, w\rangle \in \sigma(P, \langle s, w\rangle)$, and so, $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash P(y)$.

1.2 If instead $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Diamond\nabla x = y$, then $(\{\xi_s\}_{s\in S}, \langle s', w'\rangle) \vDash x = y$, for some $\langle s', w'\rangle \in Q$, and so $c(\xi_{s'}(x), \langle s', w'\rangle) = c(\xi_{s'}(y), \langle s', w'\rangle)$. Since c is 1-1, then $\xi_{s'}(x) = \xi_{s'}(y)$. Because $P(x)$ is atomic, $\xi_s(x)$ is defined. It follows that $\xi_s(y)$ is also defined and $\xi_s(x) = \xi_s(y)$. Thus, $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash x = y$. By (1.1), $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash P(y)$.

2. Let $\phi(x) = \Box\psi(x)$. Assuming $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Box\psi(x)$, take any $\langle s, w'\rangle \in Q$.

2.1 If $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash x = y$, then $\xi_s(x) = \xi_s(y)$ and, so, $(\{\xi_s\}_{s\in S}, \langle s, w'\rangle) \vDash x = y$. Since $(\{\xi_s\}_{s\in S}, \langle s, w'\rangle) \vDash \psi(x)$, by inductive hypothesis $(\{\xi_s\}_{s\in S}, \langle s, w'\rangle) \vDash \psi(y)$, thus $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Box\psi(y)$.

2.2 If instead $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Diamond\nabla x = y$, then $(\{\xi_s\}_{s\in S}, \langle s', w''\rangle) \vDash x = y$ for some $\langle s', w''\rangle \in Q$. Thus, $(\{\xi_s\}_{s\in S}, \langle s, w'\rangle) \vDash \Diamond\nabla x = y$. By inductive hypothesis, $(\{\xi_s\}_{s\in S}, \langle s, w'\rangle) \vDash \psi(y)$, and therefore $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Box\psi(y)$.

3. Let $\phi(x) = \Delta\psi(x)$. Assuming $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Delta\psi(x)$, take any $\langle s', w\rangle \in Q$.

3.1 If $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash x = y$, then $(\{\xi_s\}_{s\in S}, \langle s', w\rangle) \vDash \nabla x = y$ and so, trivially, $(\{\xi_s\}_{s\in S}, \langle s', w\rangle) \vDash \Diamond\nabla x = y$. Since $(\{\xi_s\}_{s\in S}, \langle s', w\rangle) \vDash \psi(x)$, by inductive hypothesis $(\{\xi_s\}_{s\in S}, \langle s', w\rangle) \vDash \psi(y)$. Thus, $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Delta\psi(y)$

3.2 If instead $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Diamond\nabla x = y$, then $(\{\xi_s\}_{s\in S}, \langle s'', w'\rangle) \vDash x = y$, for some $\langle s'', w'\rangle \in Q$. Hence, $\xi_{s''}(x) = \xi_{s''}(y)$ and, so, $(\{\xi_s\}_{s\in S}, \langle s'', @\rangle) \vDash x = y$. Consequently, $(\{\xi_s\}_{s\in S}, \langle s', @\rangle) \vDash \nabla x = y$ and, so, $(\{\xi_s\}_{s\in S}, \langle s', w\rangle) \vDash \Diamond\nabla x = y$. By inductive hypothesis, $(\{\xi_s\}_{s\in S}, \langle s', w\rangle) \vDash \psi(y)$, and therefore $(\{\xi_s\}_{s\in S}, \langle s, w\rangle) \vDash \Delta\psi(y)$.

Q.E.D.

Finally, it is worth remarking that SC-validity is not preserved under uniform substitution. For instance, $P(x) \to x = x$ is SC-valid, whereas $\sim P(x) \to x = x$ is not.

17.2.2.2 $\mathcal{L}_\Box$-Logic

The next problem is determining which typical laws and inference rules of $\mathcal{L}_\Box$ carry over to $\mathcal{L}_{\Box\Delta}$. First of all, it is noteworthy that the *rule of necessitation* fails in $\mathcal{L}_{\Box\Delta}$, since $\nabla x = x$ is valid, whereas $\Box\nabla x = x$ is not. The same rule however is SC-valid in the sub-language $\mathcal{L}_\Box$:

N$^-$. If $\vDash \phi$, then $\vDash \Box\phi$, for $\phi \in \mathcal{L}_\Box$

Proof. Choose a model $\mathcal{M}$ and an assignment $\{\xi_s\}_{s\in S}$. Given any precisification s_0, pick out a world $\langle s_0, w_0\rangle \in Q_{\mathcal{M}}$. Now, consider the one-precisification model $\mathcal{M}'$ which is obtained by restricting $\mathcal{M}$ to s_0, and let $@_{\mathcal{M}'} = w_0$. Define in $\mathcal{M}'$ the assignment $\xi'_{s_0}(x) = c(\xi_{s_0}(x), \langle s_0, w_0\rangle)$. Since $\vDash \phi$, then $(\mathcal{M}', \xi'_{s_0}, \langle s_0, w_0\rangle\}) \vDash \phi$. Since $\phi \in \mathcal{L}_\Box$, the truth of ϕ at a world is independent of what is the case at any other world from a different precisification. So, $(\mathcal{M}, \{\xi_s\}_{s\in S}, \langle s_0, w_0\rangle\}) \vDash \phi$. Hence, $(\mathcal{M}, \{\xi_s\}_{s\in S}, \langle s_0, @_{\mathcal{M}}\rangle\}) \vDash \phi$. Q.E.D.

On the other hand, it is easy to show that the *Kripke axiom*

K. $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$

is SC-valid in $\mathcal{L}_{\Box\Delta}$, unlike in some counterpart-theoretic frameworks (most notably, the one in Lewis [9]).

The following major modal theses are all SC-valid:

T. $\Box\phi \to \phi$
B. $\phi \to \Box\Diamond\phi$
4. $\Box\phi \to \Box\Box\phi$
5. $\Diamond\phi \to \Box\Diamond\phi$

This is another respect in which the present semantics differs from Lewis' counterpart theory. For in the latter (and restrictedly to $\mathcal{L}_\Box$), these four theses hold only if the counterpart relation is reflexive, symmetric, transitive and euclidean, respectively. In SC semantics, on the other hand, we need not make such assumptions concerning counterparthood.

It is also easy to show that the *Barcan schema* and its *converse* hold in $\mathcal{L}_{\Box\Delta}$:

BF. $\Diamond\exists x\phi \to \exists x\Diamond\phi$
CBF. $\exists x\Diamond\phi \to \Diamond\exists x\phi$

Let us take a look now to the modal properties of identity. The *necessity of identity* and *non-identity* are both SC-valid:

NI. $x = y \to \Box x = y$
NN. $\sim x = y \to \Box{\sim} x = y$

The *necessity of self-identity* $\Box x = x$, on the other hand, fails (which follows immediately from **T** and the invalidity of $x = x$). It follows that **NI** can't be proved in the usual way from the conjunction of **LL** and the necessity of self-identity. Nevertheless, the following weakened versions hold:

NSI_1^-. $\forall x\Box x = x$
NSI_2^-. $x = x \to \Box x = x$

The following are also SC-valid theses:

NSD. $\sim x = x \to \Box{\sim} x = x$
NE. $\exists y(x = y) \to \Box\exists y(x = y)$
NNE. $\sim\exists y(x = y) \to \Box{\sim}\exists y(x = y)$

The following four valid SC-schemas show how blocks of modal operators can be simplified to a single modal operator:

$\Box\Diamond$. $\Box\Diamond\phi \leftrightarrow \Diamond\phi$
$\Diamond\Diamond$. $\Diamond\Diamond\phi \leftrightarrow \Diamond\phi$
$\Diamond\Box$. $\Diamond\Box\phi \leftrightarrow \Box\phi$
$\Box\Box$. $\Box\Box\phi \leftrightarrow \Box\phi$

Proof. ($\Box\Diamond$) by (**T**), (**5**). ($\Diamond\Diamond$) by (**T**), (**4**). ($\Diamond\Box$) by ($\Box\Diamond$). ($\Box\Box$) by ($\Diamond\Diamond$).

17.2.2.3 $\mathcal{L}_\Delta$-Logic

The topic of this subsection are the laws and rules of $\mathcal{L}_\Delta$ which are SC-valid in $\mathcal{L}_{\Box\Delta}$.

The rule of Δ-introduction, typical of supervaluationism, holds:

ΔI. $\phi \vDash \Delta\phi$

From Δ**I** it follows that

ΔN. If $\vDash \phi$, then $\vDash \Delta\phi$

the analog of necessitation, which guarantees that valid formulas are closed under definiteness. The analog of the Kripke axiom is SC-valid, too:

ΔK. $\Delta(\phi \to \psi) \to (\Delta\phi \to \Delta\psi)$

Insofar as we are presupposing that admissibility is absolute, the following are all SC-valid:

ΔT. $\Delta\phi \to \phi$
ΔB. $\phi \to \Delta\nabla\phi$
Δ4. $\Delta\phi \to \Delta\Delta\phi$
Δ5. $\nabla\phi \to \Delta\nabla\phi$

Since world domains can vary across precisifications, the analog of the Barcan Schema, $\nabla\exists x\phi \to \exists x\nabla\phi$ fails. So does its converse $\exists x\nabla\phi \to \nabla\exists x\phi$, for instance when ϕ is $\sim x = x$.

Operators for semantic (in)determinacy can be simplified as follows:

$\Delta\nabla$. $\Delta\nabla\phi \leftrightarrow \nabla\phi$
$\nabla\nabla$. $\nabla\nabla\phi \leftrightarrow \nabla\phi$
$\nabla\Delta$. $\nabla\Delta\phi \leftrightarrow \Delta\phi$
$\Delta\Delta$. $\Delta\Delta\phi \leftrightarrow \Delta\phi$

Analogously to the modal case, the proof employs a combination of (Δ**T**), (Δ**4**) and (Δ**5**). Moreover, ($\Delta\nabla$) and ($\nabla\nabla$) entail, respectively,

$\Delta\nabla^*$. $\Delta I\phi \leftrightarrow I\phi$
$\nabla\nabla^*$. $\nabla I\phi \leftrightarrow I\phi$

In particular, $(\Delta\nabla^*)$ rules out the possibility of higher-order vagueness. As mentioned in Sect. 17.1, however, unrestricted quantification cannot be definite at any order, i.e., it can be vague only if it is super-vague. Therefore, as long as we want to capture vague quantification over worlds or world-bound individuals, the SC-semantics will need to be relaxed by introducing a suitable admissibility relation, thus obtaining a weaker logic of definiteness—arguably one in which (Δ**4**) and (Δ**5**) fail. I leave such refinements for another time.

17.2.2.4 $\mathcal{L}_{\Box\Delta}$-Logic

This subsection is devoted to a number of conditions on the interaction of modal and determinacy operators. We will then proceed to determine which ones are SC-valid.

Although the literature does not offer any specific work on the combination of modal and supervaluationary logic, there is a good deal of work on *product logics* for multi-modal languages. A product logic is defined semantically with respect to a class of models that are the cartesian products of Kripke models.[19] Product logics validate three key principles whose analogs in $\mathcal{L}_{\Box\Delta}$ are:

Commutativity$_1$. $\Diamond\nabla\phi \to \nabla\Diamond\phi$
Commutativity$_2$. $\nabla\Diamond\phi \to \Diamond\nabla\phi$
Church-Rosser. $\Diamond\Delta\phi \to \Delta\Diamond\phi$

Are these schemas SC-valid? Consider the following conditions on a SC-frame $\mathcal{F}$, for $s, s' \in S$, $w \in W$:

C$_1$. If $\langle s, @\rangle, \langle s', @\rangle, \langle s', w\rangle \in Q$, then $\langle s, w\rangle \in Q$
C$_2$. If $\langle s, @\rangle, \langle s, w\rangle, \langle s', w\rangle \in Q$, then $\langle s', @\rangle \in Q$
CR. If $\langle s, @\rangle, \langle s', @\rangle, \langle s, w\rangle \in Q$, then $\langle s', w\rangle \in Q$

It it is not hard to check that **Commutativity**$_1$ (**Commutativity**$_2$**, Church-Rosser**) is true in every model based on a frame $\mathcal{F}$ iff **C**$_1$ (**C**$_2$**, CR**) holds in $\mathcal{F}$. Now, the consequent of **C**$_2$ is trivially satisfied in every SC-frame, since $\langle t, @\rangle \in Q$, for every $t \in S$. It follows that **Commutativity**$_2$ is SC-valid. Notice, however, that the necessitation of **Commutativity**$_2$, $\Box(\nabla\Diamond\phi \to \Diamond\nabla\phi)$, is invalid.[20] This is one of those cases in which the rule of necessitation fails in $\mathcal{L}_{\Box\Delta}$. On the other hand, neither **C**$_1$ nor **CR** are true of every SC-frame, hence both **Commutativity**$_1$ and **Church-Rosser** are invalid.

Now, call a SC-frame *complete* when $Q = S \times W$, i.e., when the set of all worlds contains no gaps across precisifications. It should be clear that **Commutativity**$_1$ and **Church-Rosser** (and, trivially, **Commutativity**$_2$) are all valid with respect to the class of complete SC-frames. The moral is that those three conditions hold when

[19] See, for instance, Gabbay et al. [4], ch. 5.

[20] Because **C**$_2$ is no longer trivially true, but in fact can be false, when @ is replaced with an arbitrary $u \in W$.

the set of worlds is determinate. In fact, the following schemas are also valid with respect to the complete SC-frames:

$\Box$-Commutativity$_1$. $\Box(\Diamond\nabla\phi \rightarrow \nabla\Diamond\phi)$
$\Box$-Commutativity$_2$. $\Box(\nabla\Diamond\phi \rightarrow \Diamond\nabla\phi)$
$\Box$-Church-Rosser. $\Box(\Diamond\Delta\phi \rightarrow \Delta\Diamond\phi)$

I now turn to one of the most interesting conditions concerning the logic of $\mathcal{L}_{\Box\Delta}$, which is

Locality. $I\Diamond\phi \rightarrow \Diamond I\phi$

This schema captures the idea that any instance of indeterminacy about what is the case over the whole logical space reduces to an instance of indeterminacy at some particular world. If the schema is invalid, we say that modal vagueness can be *global*.

As it turns out, **Locality** is SC-valid with respect to the class of complete frames. On the other hand, it is easy to construct countermodels over SC-frames which are incomplete. Let W be a property uniquely instantiated by some world w. If w exists at some but not all precisifications, then

12. It is vague whether W could possibly be instantiated.

On the other hand, no world is such that it instantiates W at some but not all precisifications. So,

13. It could possibly be vague whether W is instantiated

is false. Or, to use a concrete example, recall that vagueness about what worlds there are, absolutely, can arise from issues of plenitude. Following Sect. 17.1.4, suppose that the possible size of any world has an indeterminate upper bound—let the candidate values be k_1 and k_2—, but that any given world has precise size. Then it is true that

14. It is vague whether there could possibly exist k_2 duplicates of the Tower of Pisa

since it is indeterminate whether logical space contains worlds large enough to fit k_2 objects. On the other hand, it is not the case that

15. It could possibly be vague whether there exist k_2 duplicates of the Tower of Pisa

since there is no world whose size is indeterminate.

17.3 Revisionism?

I have put forward a language $\mathcal{L}_{\Box\Delta}$ with modal and determinacy operators, whose logic is defined by a combination of counterpart-theoretic and supervaluationary semantics. In Barnes and Williams [1] it has been argued, however, that a language as rich as $\mathcal{L}_{\Box\Delta}$ will have to make some modal inconsistency satisfiable, if vagueness is interpreted via supervaluations.

Let us look at the objection in more detail. The argument in Barnes and Williams [1] is preceded by the observation that supervaluationary logic is perfectly classical with respect to an extensional language $\mathcal{L}$. Indeed, this fact is typically exhibited as a virtue of supervaluationism *vis á vis* alternative semantics for vagueness, especially those of the degree-theoretic variety. Some of that classicality gets 'lost', as it were, once the language is enriched with a determinacy operator, hence expanded to $\mathcal{L}_{\Delta}$. Indeed, in such languages, *reductio ad absurdum* and other classical forms of inference fail. The main charge of Barnes and Williams [1] is that, once we add modal operators as well and define a supervaluationary logic for $\mathcal{L}_{\Box\Delta}$, the departure from classical logic would be unacceptable insofar as some inconsistencies become satisfiable. The argument goes as follows. Given a language $\mathcal{L}_{\Box\Delta}$, take some ϕ such that

(a) $\nabla{\sim}\phi \wedge \nabla\phi$

Since $\phi \vee {\sim}\phi$ is supervaluationarily valid, we can infer

(b) $(\phi \wedge \nabla{\sim}\phi) \vee ({\sim}\phi \wedge \nabla\phi)$

But modalities are factive, hence

(c) $\Diamond((\phi \wedge \nabla{\sim}\phi) \vee ({\sim}\phi \wedge \nabla\phi))$

Now, assume the validity of the following inferential schema—let's call it *modal reductio ad absurdum*:

MR. If $\Gamma, \phi \vDash \bot$, then $\Gamma \vDash \sim \Diamond\, \phi$

Since each disjunct in (b) is supervaluationarily inconsistent, by (**MR**) we can derive

(d) $\sim \Diamond\, (\phi \wedge \nabla{\sim}\phi)$
(e) $\sim \Diamond\, ({\sim}\phi \wedge \nabla\phi)$

But the following modal inference is clearly valid:

MD. $\Diamond(\phi \vee \psi) \vDash \Diamond\phi \vee \Diamond\psi$

By (**MD**), (c) is inconsistent with the conjunction of (d) and (e). The moral is that any language with modal and determinacy operators whose logic is supervaluationary makes inconsistent statements satisfiable, if some statements are vague.

However, the supervaluationist does not have to accept that conclusion. The argument appeals to two modal inference forms, **MR** and **MD**. The Barnes-Williams objection tacitly assumes that, if such inference forms hold in the language $\mathcal{L}_{\Box}$, their validity should carry over to $\mathcal{L}_{\Box\Delta}$. Is that so? On the one hand, **MD** not only looks very natural, but is also SC-valid in $\mathcal{L}_{\Box\Delta}$. Therefore, we have *prima facie* reasons for accepting it. On the other hand, note that **MR** entails classical *reductio ad absurdum*, provided that modality is factive (i.e., that it satisfies **T**). But we saw that classical *reductio* fails already in $\mathcal{L}_{\Delta}$, therefore we should have only expected it to fail in the richer language $\mathcal{L}_{\Box\Delta}$. If in $\mathcal{L}_{\Box\Delta}$ we accept the failure of classical *reductio*, *afortiori* we should accept the failure of the stronger modal version **MR**. It can be concluded that a supervaluationary logic for $\mathcal{L}_{\Box\Delta}$ does not have to make inconsistencies satisfiable, *pace* Barnes and Williams.

References

1. Barnes, E., and J.R.G. Williams. 2011. A theory of metaphysical indeterminacy. In *Oxford studies in metaphysics*, vol. 6, ed. K. Bennett and D. Zimmerman. Oxford: Oxford University Press.
2. Divers, J. 2002. *Possible worlds*. London: Routledge.
3. Forrest, P., and D. Armstrong. 1984. An argument against David Lewis' theory of possible worlds. *Australasian Journal of Philosophy* 62: 164–168.
4. Gabbay, D., et al. eds. 2003. *Many-dimensional modal logics: Theory and applications*. Amsterdam: Elsevier.
5. Gibbard, A. 1975. Contingent identity. *Journal of Philosophical Logic* 4(2): 187–221.
6. Hale, B. 1996. Absolute necessities. *Noûs, 30, Supplement: Philosophical Perspectives* 10: 93–117.
7. Hazen, A. 1979. Counterpart-theoretic semantics for modal logic. *The Journal of Philosophy* 76(6): 319–338.
8. Keefe, R. 2000. *Theories of vagueness*. Cambridge: Cambridge University Press.
9. Lewis, D. 1968. Counterpart theory and quantified modal logic. *The Journal of Philosophy* 65(5): 113–126.
10. Lewis, D. 1971. Counterparts of persons and their bodies. *The Journal of Philosophy* 68(7): 203–211.
11. Lewis, D. 1986. *On the plurality of worlds*. Oxford/New York: Blackwell.
12. Nolan, D. 1996. Recombination unbound. *Philosophical Studies* 84(2/3): 239–262.
13. Nolt, J. 2006. Free logics. In *Philosophy of logic*, Handbook of the philosophy of science, ed. D. Jacquette. Amsterdam: North Holland.
14. Schaffer, J. 2010. The internal relatedness of all things. *Mind* 119(474): 341–376.
15. Segerberg, K. 1973. Two dimensional modal logic. *Journal of Philosophical Logic* 2(1): 77–96.
16. Sider, T. 2003. Against vague existence. *Philosophical Studies* 114(1): 135–146.
17. Stalnaker, R. 2003. Counterparts and identity. In *Ways a world might be*, 111–132. Oxford/New York: Oxford University Press.
18. Thomason, R. 2002. Combinations of tense and modality. In *Handbook of philosophical logic*, vol. 7, ed. D. Gabbay and F. Guenther. Dordrecht: Kluwer.
19. Torza, A. Forth. Vague existence. In *Oxford studies in metaphysics*, ed. K. Bennett and D. Zimmerman. Oxford: Oxford University Press.
20. Williamson, T. 1994. *Vagueness*. London: Routledge.
21. Williamson, T. 1999. On the structure of higher-order vagueness. *Mind* 108(429): 127–143.

Chapter 18
What's in a (Mental) Picture

Alberto Voltolini

Abstract In this paper I present several interpretations of Brentano's notion of the intentional inexistence of a mental state's intentional object, that is, what that state is about. I hold moreover that, while all the interpretations in Sects. 18.1–18.5 are wrong, the penultimate interpretation focused on in Sect. 18.6, according to which intentional inexistence amounts to the *individuation* of a mental state by means of its intentional object, is right provided that it is embedded in the fully correct interpretation given in Sect. 18.7. This is because it merely provides one of the necessary conditions for this last interpretation, in which intentional inexistence amounts to the *constitution* of a mental state by means of its intentional object. Finally, I argue that both these interpretations preserve the idea, which strikes everyone as true, that an intentional object exists in the mental state about it very much in the same way as a pictorial character exists in the picture (*qua* interpreted entity) that depicts it.

18.1 Intentional Inexistence as Location in a Mental State

In a well-known passage of his *Psychology from an Empirical Standpoint*, Franz Brentano writes:

> Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction towards an object (which is not to be understood here as meaning a thing), or immanent objectivity. [6, pp. 88–89]

The notion of *intentional inexistence* Brentano there mobilizes is notoriously unclear. What does it mean for an object to intentionally exist *in* a mental state? One might be tempted to read the preposition "in" figuring in that locution in its literally locative sense: an object is in a mental state pretty much as a bottle of beer is in the fridge. Brentano himself seems to justify this literal reading. For immediately

A. Voltolini (✉)
Dipartimento di Filosofia e Scienze dell'Educazione, Universitá degli Studi di Torino, Via S. Ottavio 20, 10124 Torino, Italy
e-mail: alberto.voltolini@unito.it

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_18

after the passage I have just quoted he goes on saying: “Every mental phenomenon includes something as object within itself” (*ib.*).[1] It is clear enough, however, that he does not mean such an inclusion as a *spatial* inclusion. For a few sentences later he says that such an inclusion is intentional: mental states are “those phenomena which contain an object *intentionally* within themselves” (*ib.*; my italics). Nor could anyone really pursue this spatial reading. The following argument is clearly invalid, for although its premises are true, its conclusion is false:

1. The smallest pebble lying in this university courtyard is *in* Turin.
2. Turin is *in* my present thought
3. Hence, the smallest pebble lying in this university courtyard is *in* my present thought.

As just stated, although the premises of this argument are true its conclusion is false. For even if I am thinking of Turin, I am certainly not thinking of the smallest pebble in Turin University’s courtyard. The argument suffers from a fallacy of equivocation: while the preposition “in” in (1) has a literal locative meaning, this is not the case in (2). So, the “in” of intentional inexistence has no locative meaning.[2]

18.2 Intentional Inexistence as Possible Nonexistence

A traditional way of interpreting the notion of intentional inexistence is to appeal to the idea of the possible nonexistence of the intentional object, the target a mental state directs upon or is about. According to this idea, an intentional object may not exist, in the sense that there may be both mental states that are about objects that actually exist, as when I am thinking of Elizabeth II Windsor, and mental states that are about objects that do not actually exist, as when I am thinking of Humbeth, the actually nonexistent offspring of Elizabeth II and Humbert II, the last king of Italy, or even when I am thinking of Twardy, the impossible wooden cannon made of steel, which unlike Humbeth not only does not actually exist, but it does not exist even possibly.[3]

[1]Notoriously, this passage also shows that Brentano held that intentionality, or aboutness, is the mark of the mental, in the sense that something is a mental state iff it is an intentional state, i.e., it possesses intentionality. Yet in this paper I will remain neutral on this issue. From now onwards, read “mental state” as merely synonymous with “mental state endowed with intentionality”

[2]Cf. also Brandl [4, p. 266].

[3]The Twardy example comes from Twardowski [33, p. 101].

However, although this is a quite popular interpretation,[4] it cannot be the correct one. In Brentano, intentional inexistence has nothing to do with nonexistence. As Brentano stresses in the above quotation, intentional inexistence has rather to do with a modality of being that *all* intentional objects, not just *prima facie* nonexistent ones, possess: *immanent objectivity*, as Brentano puts it. For all intentional objects exist insofar as they are immanent (to be sure, in a nonspatial sense[5]) to the mental states themselves that are about them.[6]

Now, independently of Brentano's true conception of intentional inexistence, the claim that all intentional objects 'inexist' is something that every theory of intentional inexistence must account for. Whatever the existential status that at least *prima facie* qualifies it (existent, nonexistent…), every intentional object is such that it must intentionally exist in the state that thinks of it.

18.3 Intentional Inexistence as Narrow-Scope Existence

Thus, intentional inexistence affects all intentional objects, not merely those *intentionalia* that would be intuitively ranked as actually nonexistent. Nevertheless, one may think that, however one interprets the notion of intentional inexistence out, the fact that we ascribe it to intentional objects is prompted by cases in which we commonsensically tend to say that the intentional object in question does not exist, as in the Humbeth and Twardy aforementioned examples.[7] In this respect, one may well note that an intentional object exists in a mental state very much like a pictorial character—a *pictum*—exists in the picture that depicts it.[8] For both pictures and mental states exhibit the intentional inexistence of their objects insofar as both are typically about objects that intuitively do not exist. This is how Gilbert Harman explains it:

> Ponce de Leon searched Florida for the Fountain of Youth. […] He was looking for something. We can therefore say that his search had an intentional object. But the thing he was looking for, the intentional object of his search, did not (and does not) exist.

---

[4]Cf. e.g. this passage from Haugeland and Dennett: "A belief can be about Paris, but a belief can also apparently be about phlogiston—and there is no phlogiston for it to be about. This curious fact, the possible non-existence of the *object* of an intentional item, may seem to be an idle puzzle, but in fact it has proven extraordinarily resistant to either solution or dismissal. Brentano called this the *intentional inexistence* of the intentional objects of mental states." [16, p. 384].

[5]As we have seen before, the spatial interpretation of intentional inexistence is utterly incorrect.

[6]As has been definitely clarified by Crane [10].

[7]According to Segal [30, pp. 283–284], these cases prompted Brentano to defend his admittedly immanentist theory of intentional objects.

[8]This idea can be traced back at least to Scruton [28, p. 205]. See also Haugeland and Dennett [16, p. 384].

> A painting of a unicorn is a painting of something; it has a certain content. But the content does not correspond to anything actual; the thing that the painting represents does not exist. The painting has an intentional content in the relevant sense of 'intentional'.
>
> Imagining or mentally picturing a unicorn is usefully compared with a painting of a unicorn. In both cases the content is not actual; the object pictured, the intentional object of the picturing, does not exist. It is only an intentional object. [15, p. 34]

This comparison between a mental state and a picture may be stressed even more if one defends the claim that a mental state *is* a picture: in an *inner* picture, i.e. a *mental* picture, its intentional object exists in the same way as in an *outer* picture. But in order to rely on such a comparison it is not necessary to defend such an admittedly controversial claim. Even if one merely *takes* a mental state *as* a picture, one can go on saying that the mind simply internalizes what a picture exhibits publicly, namely the fact that the relevant intentional object intentionally 'inexists'.

Yet of course, to simply say that an *intentionale* exists in a mental state very much like a *pictum* exists in the picture that depicts it seems just to push the problem of understanding what intentional inexistence amounts to one step further without resolving it. Unless of course, we have at our disposal a suitable interpretation of what it means for a *pictum* to exist in the picture that depicts it.

Now, it seems that there is such an interpretation to hand. Let me call it a *lingualist* interpretation. The reason I so label this interpretation is straightforward. According to it, one may see the intentional inexistence of a *pictum* to be *linguistically* rendered by locating, in the appropriate linguistic description of the situation at stake, the relevant existential quantifier within the scope of a pictorial operator. A pictorial character *PC* intentionally exists in a picture *P* iff a sentence of the kind "according to the picture *P*, there is something identical with *PC*" is true in the mere narrow scope reading of the quantifier. In short, given the truth of so-called T-biconditionals of the form "'p' is true iff p", a pictorial character *PC* intentionally exists in a picture *P* iff according to the picture *P*, there is something identical with *PC*. To stick to the example given by Harman in the above quotation, we may say that a unicorn exists in a picture iff according to that picture, there is a unicorn. If as to mental states things go in the same way, then we may also say that the intentional inexistence affecting targets of mental states is captured by locating, in the appropriate linguistic description of the situation at stake, the existential quantifier within the scope of a mental operator. An intentional object *O* intentionally exists in a mental state *M* iff a sentence of the kind "according to *M*, there is something identical with *O*" is true in the mere narrow scope reading of the quantifier. In short, an intentional object *O* intentionally exists in a mental state *M* iff according to *M*, there is something identical with *O*. For instance, a unicorn exists in a thought about it iff according to that thought, there is a unicorn.[9]

If we put it in these terms, to ascribe intentional inexistence to an intentional object is a way of saying that we are not really committed to intentional objects.

[9] Though only implicit, one may find such an interpretation in Fauconnier [12]. This interpretation has been very recently revived by Kroon [20].

For, as I said, the existential quantifier in the scope of a pictorial/mental operator is a quantifier in a narrow scope position. Unlike with a wide scope position, in narrow scope the existential quantifier carries no ontological commitment.

One might now think that the easiest way to justify lingualism is to once again appeal to the nonexistence of intentional objects both of mental states and of pictures. If someone believes that there is something that is such and such, or a picture depicts that there is such a something, and yet this very something does not exist, one cannot infer from such a believing or depicting that there is something such that one believes or depicts that it is such and such.

Certainly, it is debatable whether the nonexistence of an intentional object really blocks that inference. As to depictions, one may well hold that when they are about a particular character, there is always a character they depict, independently of whether it exists. So, from the fact that, for example, in *La Gioconda* the particular pictorial character Mona Lisa smiles, we can infer that there is something such that according to *La Gioconda* it smiles, even though it does not exist.[10] *Mutatis mutandis*, the same holds if one believes that Mona Lisa smiles.

Granted, in order for the above to hold, one has to interpret the existential quantifier in such a wide scope reading as being non existentially loaded, that is, as ranging over both entities that exist and entities that do not exist.[11] Many refrain from such an endorsement. Yet for the purpose of ruling out this justification, one does not even need to endorse such an interpretation. It is enough to stress, as we have seen before, that intentional inexistence has to affect *all* intentional objects, hence also those that we would intuitively qualify as existing. Yet, with respect to the latter objects the inference from the narrow scope reading to the wide scope reading of the existential quantifier in the relevant sentence raises no problem, both in the pictorial and in the mental cases. From the fact that in Jacques-Louis David's fresco *Napoleon Crossing the Alps* Napoleon gets to the other side of the Alps, we can infer that there was someone who according to that painting got to the other side of the Alps, that is, Napoleon himself. *Mutatis mutandis*, the same holds for the corresponding belief.[12] Clearly enough, both David's painting and such a belief are about Napoleon; in other terms, Napoleon is their intentional object.[13]

---

[10] For this point see, e.g., Hyman [17, p. 66, fn. 12].

[11] As I defended e.g. in my Voltolini [35].

[12] Cf. on this Crane [9, p. 18].

[13] As I anticipated in fn. 9, Kroon [20] holds that there is room to interpret in lingualist terms Brentano's *original* idea of intentional inexistence of an intentional object, i.e., the idea he presented in 1874, at the time of the first volume of his *Psychology*. Yet Kroon himself seems to acknowledge that when this interpretation undisputedly fits some bits of what Brentano said, namely at the time (1911) of the second volume of the *Psychology*, it can provide a suitable treatment to be welcome to Brentano himself at that time only for the cases in which the intentional object of one's thought intuitively does not exist. For when such an object intuitively exists, also for the 1911 Brentano another story must be told (cf. *ib*:383–384). Yet this amounts to saying that lingualism does not give the right account of intentional inexistence, if this feature affects *all* intentional objects, both those that (intuitively) do not exist and those that (intuitively) exist.

However, lingualists can deploy a better justification as to why intentional inexistence has to be merely read as existence in narrow scope, as they claim. As Elisabeth Anscombe originally noted, intentional objects are affected by a certain indeterminacy. While I cannot hit a man without hitting a man of a given height, I can well think of a man without thinking of a man of any particular height.[14] Once again, the same holds of pictorial characters, as Ned Block maintains: if I depict a stick figure of a man, this figure does not go into the matter of clothing at all.[15] Now, indeterminacy may well be a reason why it is not legitimate to export the existential quantifier from narrow to wide scope in the relevant sentences, hence to be committed to *intentionalia*. This is clearly the case when mental operators are at stake. If when faced with a devastated corpse, Smith forms the mere conviction that the(re is a unique) murderer (and that s/he) was cruel, one cannot infer from such a conviction that *there is someone* who Smith is thinking of that was cruel. While the first statement is true, the second statement may well be false. For, even if it turned out that there is such a murderer, there is no definite individual Smith has in mind.[16] Moreover, nothing substantial would change if Smith gave a depictive form to his conviction. From the fact that according to a picture the(re is a unique) murderer (who) was cruel it does not follow that there is someone who according to that picture is cruel.

However, first of all note that, in order for the inference from an operator-containing sentence with an existential quantifier in narrow scope to the corresponding operator-containing sentence with an existential quantifier in wide scope to fail, the indeterminacy in question must be an *ontological* one, namely, a case of *objectual vagueness*: it is indeterminate whether object O and object O' are the same entity. In such a predicament, an inference of the above kind fails outright. For instance, since *being a unique golden mountain* provides no criterion of identity for merely possible entities, it is ontologically indeterminate whether the merely possible individual that is the only golden mountain in a possible world W is the same as the merely possible individual that is the only golden mountain in another possible world W'. Consequently, from the fact that it is possible that there uniquely is a golden mountain we cannot infer that there is something that possibly is a unique golden mountain.[17] Yet Anscombe's characterization of intentional objects does not mobilize ontological indeterminacy but, rather, *epistemological* indeterminacy, a notion to be captured as follows: for any property F, whenever a subject S/a picture P represents O, it may not be the case that (s)he/it represents O as having F or as not having F. Manifestly, epistemological indeterminacy does not entail ontological indeterminacy: the fact that something is represented as indeterminate with respect to some properties does not mean that it is vague. So, the fact that intentional objects are characterized by epistemological indeterminacy does not justify the claim that

[14]Cf. Anscombe [1, p. 161].

[15]Cf. Block [3, p. 655].

[16]On this point cf. Smith and McIntyre [31, pp. 30–31].

[17]I have defended this point in my Voltolini [34].

whenever the appropriate linguistic description of a mental or a pictorial situation concerns such objects, there is a failure of inference from a narrow scope reading to a wide scope reading as regards the existential quantifier involved in that description. There is, therefore, still no justification for the idea that intentional inexistence has to be captured by such a failure.

To be sure, the aforementioned cases of Smith's having mental states or depicting pictures concerning the murderer are cases in which the quantifier exportation fails. More in general, all cases of so-called *generic pictures*, pictures whose subject is generic, are cases in point. Take as an example Turner's *Rain, Steam, and Speed* which represents some locomotive or other moving forward along a bridge on a landscape's background, are cases in point. From the fact that in Turner's painting there is a locomotive moving forward we cannot infer that there is something that according to Turner's painting moves forward. The same holds for many mental states whose content is generic as well: Diogenes' famous quest for a honest man does not amount to fact that there was someone that Diogenes was looking for.[18] Indeed, all such cases are affected by ontological indeterminacy. It is indeterminate whether the locomotive whose movement would make true Turner's painting in a certain possible world is the same as the locomotive whose movement would make true Turner's painting in another possible world. *Mutatis mutandis*, the same holds of the honest man that would satisfy Diogenes' quest in a certain possible world and of the honest man that would satisfy Diogenes' quest in another possible world. Yet as we have already seen, in many other cases affecting either pictures or mental states (e.g. the previous Napoleon example) the relevant inference from the narrow scope reading to the wide scope reading of the existential quantifier is valid. Now, it must be remembered once again that intentional inexistence is said to characterize *all* intentional objects. So even in all these other cases, the relevant intentional objects (Napoleon in our example) exist *in* the respective pictures or mental states. Since in all such cases we are allowed to infer in the relevant pictorial or intentional sentence a wide scope reading of the relevant existential quantifier from the narrow scope reading, intentional inexistence cannot be understood by appealing to the mere narrow scope reading of the existential quantifier in the relevant sentence. Hence, intentional inexistence cannot be rendered by the lingualist idea that the relevant sentence mobilizes a mere narrow scope reading of the existential quantifier.

18.4 Intentional Inexistence as Mind-Dependent Existence

At this point, in order to understand the notion of intentional inexistence it is better to look elsewhere. As I have already recalled, in the very characterization of the notion Brentano himself gives in the aforementioned quotation from the

[18] As Chisholm [8, p. 201] originally envisaged.

Psychology, intentional inexistence is “immanent objectivity”, namely, the fact that the intentional object is immanent to the mental state itself that is about it. One may naturally interpret this idea of immanence in terms of an *existential dependence* relation. According to this interpretation, whenever one thinks, one does not think of a transcendent, that is mind-independent, entity, which in many cases is what it seems to be (for instance, when one is thinking of Elizabeth II, it seems that one’s thought is directed upon the flesh-and-blood English Queen). Rather, one thinks of a *mind-dependent* entity, a thought-of entity as some might say: in other words, an entity whose existence somehow depends on the existence of the mind that thinks of it (a contemplated Queen, in our example).[19] In turn, this notion of existential dependence that allegedly accounts for intentional inexistence is traditionally rendered in *modal* terms.[20] An intentional object exists in a mental state in the sense that it modally depends on the existence of that state for its existence: *necessarily*, if the intentional object exists, then the mental state that thinks of it exists as well. In a nutshell, intentional inexistence is the modally dependent existence of an intentional object on its state.

Once again, this interpretation seems to preserve the idea that both in the case of pictures and in the case of mental states their respective objects intentionally exist in them. For one can well say not only that an intentional object exists in a mental state iff it modally depends for its existence on the existence of such a state, but also that a pictorial character intentionally exists in the picture that depicts it solely in the case in which it modally depends for its existence on the existence of such a picture. Consider again Mona Lisa. Surely enough, if *La Gioconda* had not existed, Mona Lisa would not have existed either. Therefore, Mona Lisa exists in *La Gioconda* iff the former modally depends on the latter for its existence. This is how some students of Brentano took the notion of intentional inexistence in terms of immanent objectivity to be read.[21]

Unlike the previous one, according to this interpretation it is possible to existentially quantify on intentional objects. Whenever one says that whenever one thinks or a picture depicts, there is something one is thinking or that picture is depicting. One thereby gives a wide scope reading of the quantifier occurring in that mode of saying.

This result will be preserved in all the following interpretations of intentional inexistence. The present interpretation has nonetheless a further, specific, consequence. Once the intentional inexistence of an intentional object is taken to amount to its mind-dependence, one can give a stronger reading of the indeterminacy of an intentional object than the epistemological reading that Anscombe’s account mobilizes. However, this reading is not so strong as the ontological reading of such an indeterminacy, which amounts to the vagueness of the intentional object

[19]For this, rather traditional, interpretation of Brentano’s position see e.g. Smith and McIntyre [31, pp. 47–51].

[20]For this account, see e.g. Mulligan and Smith [23].

[21]Cf. Brandl [4, p. 276].

itself. Rather, it simply says that an intentional object, *qua* immanent hence mind-dependent object, is *metaphysically* indeterminate, in the sense that it is *incomplete*: for some pair of properties *P* and its complement non-*P*, an intentional object has neither *P* nor non-*P*.[22] So, the man Anscombe is thinking of without thinking of his height is incomplete for it is neither 2m tall nor non- 2m tall. This again fits with the relevant comparison in intentional inexistence between mental states and pictures. For one can well say that pictorial characters exhibit the same kind of incompleteness. Mona Lisa, for example, is neither Florentine nor non-Florentine.

Yet if one really wants to capture the idea that an intentional object is an immanent object in terms of a dependence relation, one does not have to simply say that an intentional object is a mind-dependent entity. One has to proceed further in characterizing this relation. For there are plenty of mind-dependent entities that are hardly characterizable as immanent entities. Consider institutions, laws and nations, or social entities more generally. Clearly enough, any such entity is a mind-dependent entity: in a world of thoughtless individuals, there would be no such entity. Yet it is not an entity immanent to the mind, or minds, that thinks of it.

So, if one wants to properly account for an entity to be an *immanent* entity in terms of dependence relations, the most natural thing to say is that an immanent entity depends not only *historically*, but also *constantly*, on the very *same* depending entity. On the one hand, historical dependence accounts for a dependent entity coming into existence: in order for the dependent entity *to come into* existence, another entity must already exist. On the other hand, constant dependence accounts for a dependent entity to *persist* into existence: in order for the dependent entity to persist, another entity must exist at every moment in which the former entity exists.[23] Thus, the immanence of an entity is rendered by the fact that a dependent entity depends both historically and constantly on the very same depending entity. Thus, not only an immanent entity comes into existence in virtue of the existence of another entity, but it also ceases to exist once that further entity expires as well. In a nutshell, the depending entity is responsible both for the birth and for the death of the dependent entity.

At first sight, this account is plausible for pictorial characters. If a certain picture ceases to exist, its *pictum* ceases to exist as well. Precious paintings are protected by unbreakable glass precisely because if those paintings were destroyed, their *picta* would not survive them. Yet a moment's reflection shows that such a plausibility is debatable. Many portraits of fictional characters are about entities that, though they historically depend on such portraits, appearances notwithstanding survive the destruction of such portraits. Mickey Mouse, for one, has certainly survived Disney's first cartoon on it: even if that cartoon no longer existed, Mickey is still among us. Only a sustainer of a woodoo metaphysics,[24] moreover, would believe

[22]For this notion of incompleteness cf e.g. Castañeda [7, p. 179], Parsons [24, pp. 56, 183–184].

[23]For these definitions cf. Thomasson [32, pp. 30–31].

[24]This notion comes from Walton [40, p. 385].

that tearing a photo of a person they hate into pieces would make that very person die. In fact, immanentists of this kind about pictures are forced to implausibly maintain that a painting of a real individual does not portray such an individual, but just a depicted surrogate (a depicted Napoleon in our previous example). Furthermore, this account is certainly not plausible for intentional objects of mental states. For it makes such objects unshareable: neither different people nor one and the same thinking subject at different times may share the same intentional object in their thoughts.

At the time of the first volume of his *Psychology* (1874) Brentano would have been happy with this result. For he seems to have believed then that the intentional objects mental states are about are phenomenal entities, entities that exist only when one is experiencing them.[25] Independently of whether the ascription of such a belief to Brentano is correct,[26] however, he himself would have later rejected this way of intending immanentism. As he then said, when a picture depicts a real individual it does not depict a depicted counterpart of it. Analogously, when one is thinking about that individual one is not thinking about a thought-of surrogate of it.[27]

18.5 Intentional Inexistence as Monadic Intentionality

At this point, one may suspect that what is wrong with the previous proposal is its *relational* account; that is, the idea that an intentional object exists in a mental state insofar as there is one such object that state is in a (dependence) relation with. What if to intentionally inexist for an intentional object means rather for its mental state to possess intentionality as a *monadic* property, namely, as a certain way for that state to be modified, a way suitably rendered by an *adverbial* description of such a state? Thus, commitment to, hence relation with, an intentional object is no longer required to account for intentional inexistence. If this is the case, when Ponce de Leon thinks of the Fountain of Youth, there is no (dependent) thing such a thought enters into a relation with. Rather, his thought is simply modified 'fountainyouthily', so to speak. Likewise, if Ponce had depicted 'it'.[28]

To be sure, this adverbialist proposal has to deny that a relation between an intentional object and its mental state holds in the case not only of nonexistent, but

[25]Cf. Crane [10].

[26] Kroon [20, p. 389, fn. 15] rightly points out that in [5], Brentano said he had never defended this form of immanentism.

[27]Cf. Brentano [5, pp. 77, 95–96]. Harman has recently defended the same point. Cf. [15, p. 36].

[28]For this proposal, see Kriegel [18, 19]. For Kriegel the pictorial case is admittedly more complicated since it involves the ascription of a derivative form of intentionality to pictures. See his [19, ch. 5].

also of existent, *intentionalia*. As we have seen previously, this is hard to swallow.[29] Yet adverbialism has to face a more serious problem. At first sight, both a thought of the Fountain of Youth and a thought of Nessie (to say nothing of the corresponding pictures) share a property, namely, their *being about something.* Yet in the adverbialist reconstruction, such thoughts are merely differently modified. The first thought has the monadic property of *being-intentionally-directed-fountainyouthily*, while the second thought has the monadic property of *being-intentionally-directed-Nessiely*.[30] As the use of dashes in the description of the properties should make manifest, such properties have nothing in common.[31]

The adverbialist would deny this negative conclusion. For him/her, the two monadic properties are just two species of the same genus, two determinates of the same determinable, *being-intentionally-directed-somehow*.[32] But this reply does not grasp the problem. The point is that, intuitively, the two thoughts exactly share the same property, they do not merely come under the same genus. In this respect, their relationship is closer to the one concerning, say, someone kicking a ball and someone kicking a child, rather than to the one concerning, say, a body having a certain temperature and another body having another temperature. True enough, in the latter case the two bodies exemplify different species coming under one and the same genus, that is, *having a temperature*. Yet in the former case, the two actions in question are both instances of *kicking*, not merely species of the same genus, say *being a purposive action*; as would rather be the case if, say, it were a question of kicking something and caressing something. Any reconstruction that said that the two actions are just simply kinds of purposive actions would lose something.

All in all, therefore, ruling out the relationality of intentional mental states does not seem the right way to account for intentional inexistence. Instead, one must look for the right kind of relation.

18.6 Intentional Inexistence as Individuation

Coming back therefore to a relationalist reading of intentional inexistence, one may however wonder on behalf of Brentano whether interpreting immanentism in terms of a dependence relation captured in modal terms is sufficient. One may indeed suppose that, taken as immanent existence, the intentional inexistence of the intentional object is rendered by the fact that the object depends on the mental state that is about it not only for its existence, but also for its *individuation*. In this

[29] Kriegel [19] tries to account for this problem by allowing for a veridical mental state to also entertain a relation with the existent object it is causally connected with. Such a relation grounds the possession for that state of its monadic intentionality property.

[30] I here follow Kriegel's [18] way of putting it.

[31] On a variant of this problem see Voltolini [37, pp. 141–144].

[32] Cf. Kriegel [19].

account, an intentional object exists in the mental state that thinks of it iff the former is individuated in terms of the latter, in the sense that a certain mental state's being about a certain object affects the nature of that object: if the object were not what the mental state is about, it would be a different entity.

Probably, reading "*a* exists in *b*" as "*b* individuates *a*" is the right, strengthened, reading of Brentano's immanentism. This is because it thoroughly accounts for his phenomenalism, namely for the fact that for him the physical events mental states are about are just phenomenal events.[33] For instance if a certain sound, taken as a certain phenomenal event, were not what a certain auditory experience is about, then it would not be that very sound.

Clearly enough, however, this interpretation merely aggravates the problems that the thesis of the constant dependence of *intentionalia* gives rise to. Suppose that an intentional object became a different entity if it were not intended by a certain mental state. Then it would be even more clearly unshareable by different people or even by one and the same thinking subject at different times.

Yet no such problems arises if one reads the individuative relation that intentional inexistence allegedly picks out in the *opposite* direction. According to *this* reading, intentional objects exist in mental states in the sense that *those very states* are individuated (at least in part) in terms of them. In other terms, it belongs to the nature of a certain mental state that it is about a certain object; if it were not about that object, it would not be that mental state.[34]

Here one may say again that this sense of intentional inexistence equates mental states with pictures as to their individuation. For, if one takes a picture not merely as a certain material object in the world but as an interpreted entity, that is, as a meaningful representation, a picture may be individuated in terms of the object it is about very much like a mental state is individuated by the object it is about. Consider again *La Gioconda* not merely as a certain canvas hung on a wall in the Louvre, but as a picture *of* Mona Lisa. Clearly enough, if it were not about Mona Lisa, it would not be the same painting. Simply, one may add that while pictures are individuated

---

[33]Cf. Crane [10].

[34]This is the interpretation of intentional inexistence Crane himself apparently defends: cf. [9, p. 29]. I say "apparently" for two reasons. First of all, since Crane basically believes that there are no nonexistent *intentionalia*, a thought about a nonexistent *intentionale* is for him not relational. Hence it can at most be *identified* but not individuated in terms of its intentional object; the description of such a state as being about that object distinguishes that state from any other such states. Crane himself admits this [9, p. 31]. Indeed, in his proper theory of intentionality, a mental state is individuated by its intentional *content*: any mental state has its own intentional content independently of whether there really is the *intentionale* it is about [9, p. 32]. Moreover, in order for such an appeal to individuation to account for intentional inexistence, it must account for our saying that an intentional object is in the state that is about it, but not for the converse idea that the state is *in* the object it is about. Yet Crane sometimes appeals to intentional individuation to account for the latter idea: see [9, pp. 82–83].

in terms of the objects they are about only insofar as they are taken as interpreted entities, as entities-*cum*-meaning, mental states are instead straightforwardly individuated in such terms since they possess intentionality on their sleeves. As many say, unlike pictures their intentionality is original and not derived.[35]

A clear advantage of this interpretation is that, unlike all the previous interpretations, it really conserves the idea that intentional inexistence affects *all* intentional objects. Every intentional object intentionally exists in the mental state that is about it insofar as it contributes to the individuation of that mental state. Not only Elizabeth II individuates a thought about her but also Humbeth, his nonexistent son, individuates the corresponding thought about him.

Here one might immediately wonder: how can something that does not exist individuate a mental state, by so being in a relation with it? If an intentional object does not exist, there is no such thing, hence there is nothing that state can be in a relation of individuation with. One should not, however, be led astray here by the problem of nonexistence. As Tim Crane first pointed out, an intentional object is a *schematic* object, namely, an object that has no nature insofar as it is thought of.[36] Put in more positive terms, schematic objects are entities that have a nature which may well vary from one schematic object to another, but it is not determined by the fact that such objects are thought of.[37] So, there may be thoughts about concrete entities—for example, an actually spatiotemporal individual such as Napoleon—very much like thoughts about nonconcrete entities—for example, an utterly non-spatiotemporal entity such as the Napoleonic Code. For, although those entities have a different nature, in the thoughts that respectively think of them they are just intentional objects, mere targets of such thoughts. True enough, it may turn out that entities of a certain nature are not allowed to figure in the overall ontological domain of what there is. For instance, fictionalists about numbers reject any ontological commitment to numbers. Now, if we accept such a rejection, this simply entails that mental states apparently about such entities are about different things that are really there. As a result, they will be individuated by the latter things. In my previous example, if there are no numbers, thoughts apparently about them are rather individuated in terms of other entities such thoughts are really about and that are really there, namely certain make-believe practices. Now, insofar as nonexistent things may well belong to kinds of entities that are admitted in the overall domain of what there is, thoughts that think of them may really be individuated by them, independently of the fact that they do not exist.[38] This may clearly be the case not only with our thought of (the nonexistent) Superman, if Superman is a fictional entity and we allow for *ficta*,[39] but also with our thought of (the nonexistent) Humbeth, if Humbeth is a merely possibly spatiotemporal individual and we allow

[35]Cf. e.g. Dretske [11], Fodor [14], Searle [29].

[36]Cf. Crane [9].

[37]I read's Crane idea of schematicity in these terms in my Voltolini [39].

[38]I have defended those points in my Voltolini [37, 39].

[39]As all realists about fictional entities believe: see e.g. Thomasson [32].

for mere *possibilia*. All in all, therefore, a mental state can well stand in a relation of individuation with an intentional object that does not exist. For there really is such an object,[40] provided it belongs to a kind of thing that already belong to the overall ontological domain.[41]

In a nutshell, as well as immanentism in all its forms, the individuative conception of intentional inexistence allows for quantifying over intentional objects, independently of whether they exist. In this sense, quantification over such objects is existentially unloaded: it is merely *particular* (or partial) quantification, as some put it.[42] If an intentional object individuates the mental state about it, then *there* is something that individuates such a state. In this respect, this quantification is the mark of ontological commitment, Quine notoriously held.[43] Yet one can really quantify over intentional objects, hence be committed to them, only insofar as they belong to a kind of entity which one can independently quantify over, hence one is independently committed to. In other words, quantification over, hence commitment to, intentional objects, is parasitic on quantification over, hence commitment to, entities that have the same nature as such *intentionalia*. This result will also be preserved in the next and final interpretation of intentional inexistence.

18.7 Intentional Inexistence as Constitution

What I have just done, then, is to reverse the direction of interpretation in the relevant reading of intentional inexistence. In accounting for intentional inexistence, instead of holding the idea that intentional objects have to be individuated in terms of the mental states that think of them, I rather stick to the opposite idea that such mental states have to be individuated in terms of their intentional objects.

Yet at this point one might wonder why, in order to perform one such reversal, one really needs to read intentional inexistence as the relation of individuation of mental states in terms of intentional objects. Cannot one read intentional inexistence as the weaker relation of mere existential dependence, but now taken as a modal dependence of mental states on such objects, rather than as the converse modal dependence of intentional objects on mental states (as in the aforementioned

[40] Of course, if the case in question mobilized an impossible relation to no *relata*, it would conceptually be a very complicated case. But the case in question is simply the case of a relation to nonexistent *relata*. That there is no conceptual problem with relations to nonexistents provided that there already are such nonexistent *relata* has been recently argued also by Priest [25, p. 60, fn. 7].

[41] As I have claimed in my Voltolini [36].

[42] Cf. e.g. McGinn [22, pp. 32–37].

[43] To be sure, Quine famously took such a quantification as existentially loaded quantification. Cf. Quine [26]. For some reasons as to why this quantification must be taken as merely particular quantification (thereby severed from a first-order property of existence that only some entities possess), see my Voltolini [38].

Brentano form of immanentism)? Why does one need to say that an intentional object exists in a mental state iff the latter is individuated in terms of the former rather than simply saying that such an inexistence occurs iff there cannot be the latter without the former?[44]

The reason is straightforward: modally conceived existential dependence of the mental state on its intentional object is not fine-grained enough to account for intentional inexistence. Consider two entities which are *mutually* existentially dependent, in the sense that both the former cannot exist without the latter and the latter cannot exist without the former, as for example, Socrates and his singleton. Moreover, suppose that someone is thinking of Socrates. We would like to say that Socrates, not his singleton, exists in such a thought. Yet because of the above predicament we have that such a thought depends for its existence both on Socrates and on his singleton: necessarily, if there is such a thoughts, there are both Socrates and its singleton. [45] So, nothing weaker than individuation of the mental state in terms of its intentional object may account for the fact that the latter intentionally exists in the former.

However, one may also wonder whether something *stronger* than individuation of the mental state in terms of its object is needed to account for intentional inexistence. According to Johannes Brandl, this is precisely the case. For Brandl, intentional inexistence has to be interpreted as the idea that the intentional object is a *necessary* part of the mental state that thinks of it.[46]

This *mereological* conception of intentional inexistence is certainly stronger than the individuative conception. Moreover, one may well concede that this mereological conception is definitely *needed* in order to correctly account for intentional inexistence. For the individuative conception is not enough, as again the case of Socrates and his singleton may well show. As Kit Fine has maintained,[47] although Socrates and his singleton are mutually dependent entities, there clearly is one sense according to which Socrates is prior to his singleton. This is precisely the individuative sense. Unlike Socrates, Socrates' singleton needs Socrates for its individuation: if there were no Socrates, Socrates' singleton would not be the thing it is—but not the other way around. Yet clearly enough Socrates does not intentionally exist in his singleton since it can be argued that although in his individuative role Socrates is a member of his singleton, he is no part of it, let alone a necessary part of it.[48]

[44]The aforementioned Mulligan and Smith [23] precisely go in this direction.

[45]On this problem see Sacchi and Voltolini [27].

[46]Cf. Brandl [4, p. 274]. Clearly enough, an intentional object being merely a *contingent* part of the thought would not account for intentional inexistence. Not only would we risk falling again into the wrong locative interpretation of intentional inexistence, but also we would mobilize something definitely weaker than the relation of individuation of the mental state in terms of its object.

[47]Cf. Fine [13, pp. 271, 279].

[48]Granted, unless one reconstructs set theory in mereological terms, *à la* Lewis [21].

It is, however, hard to see this proposal as yielding an individually *sufficient* condition for intentional inexistence. For one thing, an intentional object being a necessary part of the mental state that thinks of it might even be read back in terms of the already rejected Brentanian forms of immanentism, those according to which the intentional object either modally depends on or is individuated in terms of the mental state that thinks of it. Indeed, that object might be a necessary part of that state and still modally depend on or be individuated in terms of it.

So, the natural suggestion is to take the individuative conception and the mereological conception as providing necessary and jointly sufficient conditions of intentional inexistence. An intentional object exists in the mental state that is about it iff (i) the former individuates the latter (ii) the former is a necessary part of the latter, if it exists. According to this idea, I can say that intentional inexistence is the very relation of *constitution* of the mental state in terms (at least in part) of its intentional object. For that object not only individuates the state but is also a necessary part of it, if it exists, *qua* one of its constituents. This prevents not only Socrates, who merely individuates his singleton, from intentionally existing in such a singleton, but also a thought-of Socrates, who (if there were any) could be a necessary part of the relevant thought but definitely does not individuate it, from intentionally existing in it.[49]

Once again, this way of putting things equates mental states and pictures as to their individuation. For *qua* interpreted entities, entities-*cum*-meaning, pictures have pictorial characters among their constituents: the latter both individuate and are necessary parts of the former, if such pictures exist. Also, insofar as the intentional object that exists in a mental state stands in a relation of constitution with such a state, this intentional object can again be quantified over, hence be committed to. As we have already seen, if the object individuates the state, then *there is* something that individuates the latter. Since, moreover, constitution is just a stronger relationship than individuation, whatever ontological commitment individuation involves, it is also involved by constitution.

As a final consequence of this idea, the intentional existence of the object in the mental state about it is nothing other than the converse of the property of intentionality itself, the property for that very mental state of being about its object. For intentionality itself may well be conceived as the *very* relation for that state of *being constituted* by its intentional object. This consequence would certainly fit in with Brentano's original desideratum. For in the quotation from the *Psychology* we started from, he keeps intentional inexistence and intentionality together. To quote that passage again, "every mental phenomenon is characterized by what the

[49]The clause "if the state exists" further distinguishes my proposal from Brentanian immanentism. *Qua* constituent of a mental state, an intentional object is part of it only wherever there is such a state. This allows for that object to figure also in a possible world where there is not such a state. On the contrary, Brentanian immanent objects entail the existence of the states of which they are necessary parts.

Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call [. . .] direction towards an object".[50] Granted, however, this is a topic for another story.[51]

References

1. Anscombe, G.E.M. 1965. The intentionality of sensation: A grammatical feature. In *Analytic philosophy*, ed. R. Butler, 2nd series, 158–180. Oxford: Blackwell.
2. Aquila, R.E. 1977. *Intentionality: A study of mental acts*. University Park: Pennsylvania State University Press.
3. Block. N. 1983. The photographic fallacy in the debate about mental imagery. *Noûs* 17: 651–661.
4. Brandl, J. 1996. Intentionality. In *The school of Franz Brentano*, ed. L. Albertazzi, M. Libardi, and R. Poli, 261–284. Dordrecht: Kluwer.
5. Brentano, F. 1966. *The true and the evident*. London: Routledge and Kegan Paul.
6. Brentano, F. 1995. *Psychology from an empirical standpoint*. London: Routledge.
7. Castañeda, H.N. 1989. Fiction and reality: Ontological questions about literary experience. In *Thinking, language, and experience*, 176–205. Minneapolis: University of Minnesota Press.
8. Chisholm, R.M. 1967. Intentionality. In *Encyclopedia of philosophy*, ed. P. Edwards, 201–204. London: MacMillan.
9. Crane, T. 2001. *Elements of mind*. Oxford: Oxford University Press.
10. Crane, T. 2006. Brentano's concept of intentional inexistence. In *The Austrian contribution to analytic philosophy*, ed. M. Textor, 20–35. London: Routledge.
11. Dretske, F. 1995. *Naturalizing the mind*. Cambridge, MA: MIT.
12. Fauconnier, G. 1985. *Mental spaces*. Cambridge, MA: MIT.
13. Fine, K. 1995. Ontological dependence. *Proceedings of the Aristotelian Society* 95: 269–289.
14. Fodor, J.A. 1987. *Psychosemantics*. Cambridge, MA: MIT.
15. Harman, G. 1990. The intrinsic quality of an experience. *Philosophical Perspectives* 4: 31–52.
16. Haugeland, J., and D. Dennett. 1987. Intentionality. In *The Oxford companion to the mind*, ed. R.L. Gregory, 383–387. Oxford: Oxford University Press.
17. Hyman, J. 2006. *The objective eye*. Chicago: University of Chicago Press.
18. Kriegel, U. 2007. Intentional inexistence and phenomenal intentionality. *Philosophical Perspectives* 21: 307–340.
19. Kriegel, U. 2011. *The sources of intentionality*. Oxford: Oxford University Press.
20. Kroon, F. 2013. Intentional objects, pretence, and the quasi-relational nature of mental phenomena: A new look at Brentano on intentionality. *International Journal of Philosophical Studies* 21: 377–393.
21. Lewis, D. 1991. *Parts of classes*. Oxford: Blackwell.
22. McGinn, C. 2000. *Logical properties*. Oxford: Clarendon.

[50]Some people have taken this quotation quite literally, as saying that intentional inexistence and intentionality are one and the same thing (see e.g. Aquila [2, p. 17]). Yet I prefer to say that intentionality is the converse of intentional inexistence: intentional inexistence is the property for an intentional object to constitute the mental state that thinks of it, intentionality is the property for a mental state to be constituted by its intentional object.

[51]I hold this idea along with Elisabetta Sacchi in our Sacchi and Voltolini [27]. I thank Elisabetta, Carola Barbero, Fred Kroon and Alessandro Torza for some valuable comments on a previous version of this paper.

23. Mulligan, K., and B. Smith. 1986. A relational theory of the act. *Topoi* 5: 115–130.
24. Parsons, T. 1980. *Nonexistent objects*. New Haven: Yale University Press.
25. Priest, G. 2005. *Towards non-being: The logic and metaphysics of intentionality*. Oxford: Clarendon.
26. Quine, W.V.O. 1952. On what there is. In *Semantics and the philosophy of language*, ed. L. Linsky, 189–206. Urbana: University of Illinois Press.
27. Sacchi, E., and A. Voltolini. 2012. To think is to literally have something in one's thought. *Quaestio*, 12: 395–422.
28. Scruton, R. 1971. Intensional and intentional objects. *Proceedings of the Aristotelian Society* 71: 187–207.
29. Searle, J. 1983. *Intentionality*. Cambridge: Cambridge University Press.
30. Segal, G. 2005. Intentionality. In *The Oxford handbook of contemporary philosophy*, ed. F. Jackson and M. Smith, 283–309. Oxford: Oxford University Press.
31. Smith, D.W., and R. McIntyre. 1982. *Husserl and intentionality*. Dordrecht: Reidel.
32. Thomasson, A.L. 1999. *Fiction and metaphysics*. Cambridge: Cambridge University Press.
33. Twardowski, K. 1977. *On the content and object of presentations*. The Hague: Martinus Nijhoff.
34. Voltolini, A. 2000. Are all alleged possible objects there? *Topoi* 19: 209–219.
35. Voltolini, A. 2005. How to get intentionality by language. In *Intentionality: Past and future*, ed. G. Forrai and G. Kampis, 127–141. Amsterdam: Rodopi.
36. Voltolini, A. 2007. How to allow for intentionalia in the jungle. *Russell* 27: 86–105.
37. Voltolini, A. 2009. Consequences of schematism. *Phenomenology and the Cognitive Sciences* 8: 135–150.
38. Voltolini, A. 2012. All the existences that there are. *Disputatio* 32: 361–383.
39. Voltolini, A. 2013. There are intentionalia of which it is true that such objects do not exist. *International Journal of Philosophical Studies* 21: 394–414.
40. Walton, K.L. 1990. *Mimesis as make-believe*. Cambridge, MA: Harvard University Press.

Part V
Logical Systems

Chapter 19
Cross-World Identity, Temporal Quantifiers and the Question of Tensed Contents

Tero Tulenheimo

Abstract The use of temporal constructions in many-world settings calls for an analysis of what it means to speak of the same time over a number of scenarios. I argue that instants in two worlds can be compared for identity only in terms of temporal *world lines*. Times must be construed as such world lines, as links between instants in distinct worlds. I formulate a logical framework in which temporal quantifiers range over world lines, and show that the framework allows to represent truth-conditions of a large variety of sentences involving temporal expressions. I take propositional contents to be structures $\langle \mathfrak{t}, f \rangle$ with two independent components: a temporal world line $\mathfrak{t}$ and a function f from worlds to truth-values (a proposition). My framework allows me to distinguish two modes of temporal individuation: physical and intentional. I propose to reconstruct the A-theorist's 'tensed contents' as contents whose world line component is an intentionally individuated time, and I take B-theorists to hold that only physically individuated times are needed when accounting for the semantics of temporal language. My analysis avoids the A-theorist's metaphysical conclusions: we are not committed to any ontological correlates of grammatical tenses. My view goes against the B-theory too: both intentionally and physically individuated times are needed for formulating the semantics of temporal expressions.

19.1 Identity in Modal Settings

19.1.1 World Lines

The notion of identity is notoriously problematic in modal settings, i.e., in cases in which we must speak of something remaining the same over a number of contexts. One way of approaching the problem, put forward by Hintikka [13, 14] when discussing quantified modal logic, is to understand individuals in modal settings as *world lines* creating links between objects in different possible worlds—or, more

T. Tulenheimo (✉)
STL-CNRS and Department of Philosophy, University of Lille 3, Domaine Universitaire du "Pont de Bois", B.P. 149, 59653 Villeneuve d'Ascq, France
e-mail: tero.tulenheimo@univ-lille3.fr

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_19

generally, in different contexts. (By ‘context’ I mean a possible world equipped with a number of further parameters, which serve to specify a temporal and/or spatial location within the world.) As long as we need to take into account only a single context, we may think of individuals as objects wholly present in the context: as elements of its domain. However, claims of cross-context identity must be mediated by world lines. For conceptual clarity, I refer to the denizens of contexts as *local objects*. An *individual* is a link between local objects. A local object may realize or manifest an individual, but it cannot be an individual. All cross-context talk of individuals presupposes the availability of world lines. We are quite capable of carrying out such talk: we reason about what might have happened in counterfactual circumstances to various entities we have encountered and speak of the existence over time of entities with which we are acquainted. An individual can, then, be seen as a partial function $\mathfrak{i}$ defined on contexts, assigning to each context k on which it is defined a local object $\mathfrak{i}(k)$. The object $\mathfrak{i}(k)$ is a realization of $\mathfrak{i}$ and it belongs to the domain of k, whereas the individual $\mathfrak{i}$ itself does not reside in the domain of any context. The set of contexts in which an individual is realized will be termed its *modal margin*. It should be noted—though Hintikka does not—that the described understanding of cross-context identity leads us to conclude that if k_1 and k_2 are distinct contexts, a is an object in the domain of k_1, and b is an object in the domain of k_2, then both questions ‘Is a identical to b?’ and ‘Is a numerically distinct from b?’ are meaningless (cf. [57]). What we can ask is whether there is an individual $\mathfrak{i}$ such that $a = \mathfrak{i}(k_1)$ and $b = \mathfrak{i}(k_2)$. According to this analysis, the sentence ‘there is x which is P such that possibly x is Q’ is true in k_1 iff there is an individual $\mathfrak{i}$ realized in k_1 such that $\mathfrak{i}(k_1)$ is P, and there is an alternative context k_2 in which $\mathfrak{i}$ is likewise realized, the object $\mathfrak{i}(k_2)$ being Q.

Taking world lines as a starting point of our analysis of cross-context identity by no means implies denying the importance of the problems to which critics of quantified modal logic such as Quine have called attention. However, whereas Quine takes it that many sentences of the form ‘there is x which is P such that possibly x is Q’ are meaningless since we lack the means to decide whether an entity taken as a value of x in one context is the same as an entity in an alternative context, Hintikka notes that in various cases we indeed succeed in speaking of individuals as appearing in various contexts. The claim is that in these cases the notion of cross-context identity is presupposed and that this notion is best conceptualized in terms of world lines. Individuals understood as world lines can be seen as a transcendental precondition of modal talk. World lines are what must be given in order for us to be in a position to speak of individuals in modal settings. When commenting on world lines, Hintikka himself hesitates in his interpretation, often speaking of world lines epistemically as a means of recognizing or reidentifying an individual in different circumstances. Such a way of viewing world lines waters down what I take to be the main point of interest of this approach—namely the idea that the status of individuals as world lines is what makes it meaningful to speak of one and the same individual in modal settings.

The described standpoint on individuals is in clear contrast to Saul Kripke’s and David Kaplan’s popular views. Kripke’s notion of a rigid designator and

Kaplan's notion of direct reference are based on the assumption that one and the same thing can itself lie in the domain of various contexts—this being in my view a nonsensical idea. The view I advocate resembles in some respects Theodore Sider's temporal counterpart theory [51]: local objects are context-bound, in analogy to Sider's stages of individuals which are instantaneous. While in David Lewis's modal counterpart theory individuals are world-bound, he finds the notion of world-internal identity over time unproblematic [25, p. 192], which is enough to create a difference with respect to the view I defend. In a counterpart-theoretic approach, similarity comparisons between members of domains of distinct contexts—whether these contexts are instants as in Sider or temporally extended worlds as in Lewis—are seen as giving rise to counterpart relations, which enable us to make sense of quantification into modal contexts: 'there is x which is P such that possibly x is Q' means that some actual individual which is P has a counterpart in an alternative context and that counterpart is Q. The cross-context links created by counterpart relations are based on context-internal qualitative considerations: object b of k_2 is a counterpart of object a of k_1 given that a has internally to k_1 certain properties and b has internally to k_2 certain properties, and these two sets of properties have sufficiently much in common so that a and b can thereby be considered similar. Lewis [24] defends Humean supervenience—i.e., the position according to which any difference between two worlds is grounded in differences in their local qualities and that internally to a world, we obtain an exhaustive description of a temporally extended world as a whole by conjoining local descriptions that specify what goes on at each specific instant. If one accepts Humean supervenience, counterpart relations must be construed as supervening on context-internal properties. By contrast, when individuals are understood as world lines, they constitute an independent dimension in our conceptualization of many-world settings. They are cross-context links irreducible to the context-internal properties of their context-bound manifestations. Even if one describes contexts in the minutest detail, one will not have even touched the question of which world lines are defined over those contexts. The view I defend is, then, totally at odds with the spirit of Humean supervenience.

It is worthwhile to spend some words on what world lines are *not*. First, they are not individual concepts or Fregean senses (modes of presentation), i.e., functions assigning to every scenario an individual as the referent of a fixed singular term. For one thing, world lines are not language-relative and for another, they are not functions assigning individuals to worlds. Instead, their values are local objects. In particular, there is no reason whatsoever to think that one and the same description—say, 'the morning star'—would uniformly apply to each and every realization of a world line. Second, as already remarked, world lines are not criteria of identity in the epistemic sense. Their role is not to provide us means to recognize an individual in different circumstances. Third, world lines are not *essential properties* of the kind that Chisholm [5, pp. 5–7] or Quine [42, pp. 155–156] take as the last resort for those who wish to defend the meaningfulness of the idea of cross-world identity. Considering an individual $\mathfrak{i}$ as a world line by no means requires that there be a property P such that for every context k, the realization of $\mathfrak{i}$ in k has the property

P—let alone that this realization is the unique object having *P* in *k*. A fortiori, world lines are not *essences* in the sense of being properties unique to an individual in all contexts. If one wishes to find an analogue to world lines in the realm of essences, the point of comparison would be Alvin Plantinga's notion of essence [39].[1] Essences in Plantinga's sense consist of suitable *context-relative* properties. If *E* is such an essence, *E* associates every relevant context *k* with a property P_k, but this does not mean that there is a context-independent property *P* that *E* associates with every context.

Understood metaphysically, world lines give rise to a view of individuals which generalizes the position known as four-dimensionalism. Four-dimensionalism is a view about individuals persisting through time within one and the same world, according to which these individuals are not wholly present at any moment at which they exist; they have temporal parts and those parts are what we encounter at specific instants (see Lewis [25], Hawley [12]). If attention is confined to the temporal unfolding of the actual world, individuals in our sense behave as individuals in the four-dimensionalists' sense, and realizations of individuals amount to what the four-dimensionalists have chosen to call temporal parts of individuals. My notion of individual is more general than that of the four-dimensionalists: modal behavior must also be inbuilt within the individual itself, not just its temporal behavior. Individuals in my sense are not primarily conceived metaphysically. Rather, it is a precondition for any temporal and modal cognition pertaining to individuals that they be conceptualized as world lines. By being located in a specific context, we are directly confronted with local objects of that context. However, reasoning about an individual requires a certain conceptual command not only on the behavior of the individual in the specific context in which we find ourselves, but also on its temporal and modal behavior. Epistemologically, world lines have primacy over local objects. Different interpretations of their metaphysical status remain possible.

People of different philosophical convictions tend to find the notion of possible world problematic. I operate with possible worlds understood as mutually incompatible but intrinsically possible alternative scenarios. Independently of the stance we may adopt vis-à-vis the metaphysical status of possible worlds, we cannot help resorting to alternative, mutually incompatible scenarios when phrasing the semantics of modal notions—any more than we can help resorting to a domain of individuals when phrasing the semantics of first-order quantifiers. There are surely many legitimate questions to be posed concerning the nature of possible worlds. However, difficulties in answering these questions should not affect our semantic theorizing—unless we are prepared to limit ourselves to a purely extensional language, thereby leaving crucial fragments of actual language-use unaccounted for. Possible worlds can, for example, hardly be located in space-time, and if quantifying over them commits us to their existence, we appear to be committed to the existence of entities of a somewhat dubious variety. The semantic task is to provide an analysis of modal expressions. People engaged in a discourse involving

[1] This connection was pointed out to me by Manuel Rebuschi. See also [45, p. 114].

expressions for modalities may thereby be committed to entities that are dubious to some philosophers and ontologically harmless to others. Either way, the semantics of these expressions must be clarified. Afterwards it is up to the language users to decide in which discourses they wish to engage themselves, and for metaphysicians and perhaps epistemologists to assess what the actual commitments of the language users are in so doing.

19.1.2 Modes of Individuation

My analysis allows the distinguishing of two modes of individuation, to be referred to as the *physical mode* and the *intentional mode*. This is a systematization of a distinction Hintikka has drawn in connection with the semantics of modal notions by speaking of *public* and *perspectival methods of cross-identification* [13–15]. Hintikka's understanding of this distinction oscillates between the view of world lines as providing transcendental preconditions of quantification in modal settings and the view in which they are a means of recognizing an individual in different situations. According to the latter view, individuals are not world lines, they are merely reidentified by their means. In my classification, physically individuated individuals are individuals whose cross-context behavior is objectively determined by physical regularities; they are not conditioned by any agent. Prime examples of such individuals are material objects. Material objects are typically not only temporally but also modally extended in the sense that within certain limits, we may reasonably speak of what might have happened to them or in what ways they could interact with other material objects.

Hintikka considers visual perception as a paradigmatic example of what he calls the perspectival method of identification. This method gives rise to world lines on the basis of the visual information available to agents due to their spatio-temporal location. Construed in terms of reidentification, these world lines would reflect the agents' quite possibly erroneous views of how their objects of perception could be recognized in different counterfactual circumstances. Instead of such an epistemic construal, what interests me is the 'transcendental interpretation' of world lines. Among the different ways in which Hintikka approaches perspectival identification, there is also one that motivates this way of construing the idea.

When Alice sees a tomato, she is immediately confronted with the surface of half a tomato. It may be tempting to conclude that in this case the object of her perception is this surface (this would amount to holding a variant of the sense-datum theory of perception), but this conclusion can well be contested. If objects of perception were such 'impressions', an ordinary physical object could never itself be an object of perception. However, if we take impressions to be merely the epistemologically most privileged aspect of objects of perception, the conclusion does not follow. An alternative is that the objects of perceptual experience are always 'entire' individuals—whether the experience is veridical or not. Mere impressions of such objects are not entities structured enough to function as phenomenological objects; an impression dramatically underdetermines the object of perception. This is the view of perception that Hintikka [14, Ch. 10] defends. What goes on in perceptual individuation, thus understood, is that impressions in our visual field trigger entire intentional objects of perceptual experience: world lines defined over a whole collection V of contexts. We may take the contexts to be triples (w, t, l), where w is a world, t is an instant in the temporal order of w, and l is a spatial location of the relevant agent in w at t. This leads to a variant of the intentionalist theory of perception. The tomato as an object of Alice's visual perception is a world line j whose modal margin V consists of all perceptually relevant contexts.[2] The realization $\mathrm{j}(w, t, l)$ of world line j in context (w, t, l) is an 'impression', a region in Alice's visual field in w at t when Alice is located at l. With the possible exception of the context k_0 in which Alice finds herself, all contexts in V are counterfactual. (It is a necessary condition for the veridicality of Alice's perceptual experience that k_0 belongs to V.)

In fact *all* intentional objects are individuated in a way analogous to that of the objects of perceptual experience: they are relative to an agent and they are associated by the agent with a cross-context behavior over a set of contexts. I refer to world lines conditioned by an agent as *intentionally individuated* individuals. Such a world line may—but need not—coincide or partially overlap with a physically individuated individual. Their cross-world and cross-temporal behavior is determined by a complex network of agent-dependent factors: affects, past experiences, knowledge acquired, personal convictions, acknowledged ignorance, cultural background, etc. They are severely conditioned by an agent, and yet the agent cannot choose them at will.

Two factors contribute to the size of the modal margin of an intentionally individuated individual. On the one hand, to the extent the individual is temporally extended (as is typically the case with objects of thought), has capacities and dispositions, or has such-and-such appearance from such-and-such angles, an intentional object as a world line must be defined on a number of contexts. This sort of multiplicity of the relevant contexts manifests itself with physical individuals as well. The modal margin of an intentional object also tends to be expanded for quite a different reason. This second source of multiplicity is actually due to the very same reason why the analysis of perception or belief requires us to take into account a number of alternative scenarios. These attitudes or experiences do not fix the world uniquely but leave open a number of alternatives. Similarly, as G. E. M. Anscombe [1] stressed, intentional objects typically manifest *indeterminacy*. For example, a tomato as an intentional object does not have a definite size, definite weight, or definite shade of redness. When Alice thinks of a tomato, the intentional

[2]First, V contains all contexts (w, t, l) such that for all that Alice can tell on the basis of her perceptual experience, the current context could be (w, t, l). Second, for any of those contexts (w, t, l), V contains a number of further contexts (w, t, l') with locations l' nearby l, taking into account how Alice *could* have been positioned when attempting to get a better grasp of her surroundings.

object j of her thought may have realizations differing in weight: its realization $j(k_1)$ in context k_1 may weigh 3 ounces, while the realization $j(k_2)$ weighs 3.5 ounces, and $j(k_3)$ weighs 4 ounces. Such indeterminacy is an almost inevitable consequence of the fact that perception or belief leaves open a host of alternatives as to how the world could be. No single *physical* individual could behave in that way, but an intentionally individuated individual can and typically does.[3]

19.1.3 Plan of the Paper

There is no more reason to regard the notion of 'same time' as unproblematic in modal settings than there is to so view the notion of 'same individual'. In cross-contextual settings times, too, must be understood as world lines. Whenever one and the same time must be considered in distinct scenarios, it must be seen as a link between its local realizations. I develop a semantic theory for cases in which grammatical tenses and/or temporal indexicals are combined with expressions for modalities. The implications of this analysis for debates in the philosophy of time will be taken up.

In Sect. 19.2, the notion of temporal world line is developed and two modes of temporal individuation are discerned: the physical and the intentional. A formal semantics for a propositional tense logic is provided in Sect. 19.3, with temporal quantifiers (tense operators) ranging over temporal world lines. In Sect. 19.4 I discuss different phenomena that could be seen as motivating the so-called tensed theory of time (the A-theory), according to which grammatical tenses have as their ontological correlates special properties such as *futurity*, *pastness* and *presentness*. The A-theorists claim that these properties give rise to special kinds of propositional contents (tensed contents). In Sect. 19.5, I propose a reconstruction of the A-theorists' position which accords a semantically special role to many of the sentences to which the A-theorists wish to accord such a role, yet without the ontological repercussions that ensue from the A-theory. I base my reconstruction of the A-theorists' position on the distinction between physically and intentionally individuated times. The A-theorists' tensed contents are reconstructed as contents involving intentionally individuated times. What the A-theorists misconstrue as 'ontological tenses' emerge through the ways in which agents represent to themselves the modal behavior of past, present or future times.

[3]Objects of perceptual experience are a special case of intentionally individuated individuals. In connection with them, there is a third factor contributing to the size of the relevant modal margin: a large variety of counterfactual spatial perspectives must be taken into account; cf. Footnote 2.

19.2 Times in Modal Settings

I take seriously the idea that if we wish to speak of different eventualities with respect to the same time, *times* must be understood as world lines defined over alternative scenarios. We need temporal world lines to conceptualize the cross-world notion of *simultaneity*.[4] I take formal counterparts of grammatical tenses to be temporal quantifiers whose values are temporal world lines. They range, then, over function-like entities, whereas modal operators range over self-subsistent entities (worlds). Yet temporal quantifiers resemble modal operators in that they do not involve syntactic variables capable of binding occurrences of syntactically manifest variables.

19.2.1 Temporal World Lines and Natural Languages

Evaluating temporal discourse may well lead us to consider one and the same time in distinct scenarios. This can happen as soon as grammatically tensed verb forms and/or temporal indexicals are combined with constructions expressing a modality: constructions whose semantics requires considering multiple scenarios. Let us think of the following sentences:

1. (a) Jane could be in Alaska now
   (b) Mary thinks that Mrs. Brown is not at home
   (c) Jane believes that John was in Paris yesterday.

Suppose we are evaluating these sentences relative to w_0 at a certain time. Sentence (1a) invites us to consider Jane's being in Alaska in a counterfactual scenario v—when? By the semantics of 'now', the answer is clear: the time of Jane's being in Alaska in v must be the *same* as the time of evaluation of (1a). We must consider a certain time in both scenarios: w_0 and v. Sentence (1b) shows that this phenomenon is not triggered exclusively by indexicals but appears already with present tense: what (1b) affirms is that for any scenario v compatible with what Mary believes, Mrs. Brown is not at her house in v—at the very *same* time at which (1b) is evaluated. According to (1c), in every scenario v compatible with all that Jane believes, John is in Paris during the day immediately preceding the day to which the time of evaluation of (1c) belongs. In these scenarios v, we must consider times

[4]The way in which I problematize the notions of 'same time' and 'simultaneity' is not related to the relativity of simultaneity postulated by the special theory of relativity. The physical theory problematizes the 'intra-world' notion of simultaneity by laying it down that the question of whether two spatially separated events (within one and the same world) occur at the same time depends on the observer's frame of reference. What I claim is that in order to speak meaningfully of the same time in *distinct possible worlds*, times must be conceptualized as world lines. In a context in which the theory of relativity must be taken into account, we would need temporal world lines to reason about the modal properties of a fixed reference frame.

earlier than the very *same* time in relation to which (1c) is evaluated. The contents of all three sentences presuppose the possibility of viewing one and the same time in relation to more than one world. We must consider temporal world lines.

We need a convenient terminology to mark the temporal analogue of the distinction between *local object* vs *individual* made in Sect. 19.1. I opt for using the term 'instant' of temporal counterparts of objects and reserve the term 'time' for temporal counterparts of individuals. For each world w we associate a temporal domain temp(w) which is a certain set of instants. It is a part of my semantic theory that questions of the form 'Is t numerically identical to s?' and 'Is t numerically distinct from s?' are meaningless if t and s are instants belonging to the temporal domains of distinct worlds. In order to approach time-related questions of cross-world identity, we need to ask whether there is a temporal world line linking instants of distinct worlds. I take times to be such temporal world lines. Times are, then, partial functions taking worlds as their arguments and yielding instants as their values: if $\mathfrak{t}$ is a time realized in w, then $\mathfrak{t}(w)$ is an instant in the temporal domain of w. Instants $t \in \text{temp}(w)$ and $s \in \text{temp}(v)$ are realizations of the same time $\mathfrak{t}$, if $\mathfrak{t}$ is realized in both worlds w and v, and satisfies: $t = \mathfrak{t}(w)$ and $s = \mathfrak{t}(v)$.

The distinction instant/time forces the following question upon us: are those sentences that are semantically dependent on a temporal context to be evaluated relative to instants or times? Since we evaluate sentences relative to worlds, and as the temporal domain of a world normally contains many instants, the evaluation being temporally determinate requires that a specific instant be fixed. On the other hand, a single instant need not reveal a unique temporal world line. The same instant s_0 can be the realization of two or more times. This possibility cannot be excluded if at least one of the times considered is intentionally individuated. It might even happen that there is no time $\mathfrak{t}_0$ with $s_0 = \mathfrak{t}_0(w_0)$. In connection with modal sentences, we need to be given an entire time before the evaluation may proceed. Generally, sentences are evaluated relative to worlds and times. The semantics of *atomic* formulas will however depend only on the instants realizing times in worlds.

According to the semantics I develop in Sect. 19.3, the truth-conditions of the sentences (1a–c) can be respectively expressed as follows, when evaluated in world w_0 at time $\mathfrak{t}_0$:

2. (a) There is a scenario v such that Jane is in Alaska in v at instant $\mathfrak{t}_0(v)$.
   (b) For every scenario v compatible with all that Mary believes in w_0 at instant $\mathfrak{t}_0(w_0)$, Mrs. Brown fails to be at her house in v at instant $\mathfrak{t}_0(v)$.
   (c) For every scenario v compatible with everything Jane believes in w_0 at instant $\mathfrak{t}_0(w_0)$ and for each time $\mathfrak{t}_1$ for which $\mathfrak{t}_1(v)$ is an instant of the day immediately preceding the day to which the instant $\mathfrak{t}_0(v)$ belongs, John is in Paris in v at instant $\mathfrak{t}_1(v)$.

The need for postulating temporal world lines is not restricted to grammatically tensed sentences. Suppose for the sake of discussion that there are tenseless verb forms that can be used for expressing timelessly prevailing relationships. Let us consider evaluating in w_0 the sentence

3. On May 7, 1906, Mary *thinks* that Mrs. Brown *is* not at home on May 7, 1906,

with 'thinks' and 'is' in italics marking a tenseless use of the corresponding verbs.[5] Sentence (3) expresses that Mary's thinking takes place on a certain definite date and that according to what Mary then thinks, Mrs. Brown is not at home then. The date must be considered in relation to the set V of all scenarios compatible with what Mary believes in w_0 on May 7, 1906. The truth of (3) requires that there be a time $\mathfrak{t}_0$ such that (2b) holds: its truth-condition employs a temporal world line. Furthermore, by making use of date expressions (3) serves to locate the time $\mathfrak{t}_0$ within an agent-independent time scale. Supposing we take date expressions to denote instants (instead of times), this means that $\mathfrak{t}_0(v) =$ the denotation of 'May 7, 1906' in v, whenever $v \in V \cup \{w_0\}$.[6]

The examples analyzed thus far show that contents of certain temporal sentences involving expressions for modalities utilize cross-world links between instants.[7] Taking times to be such links allows generalizations that would not be available if we operated with point-like instants instead of temporal world lines. Two modes of individuation of individuals were distinguished in Sect. 19.1.2: the physical and the intentional. In connection with semantic values of temporal constructions, a similar distinction is called for. I will explore to which extent and how we can utilize the distinction between intentional and physical *temporal* world lines when analyzing the contrast between tensed and tenseless sentences. In Sects. 19.4 and 19.5 I will shed new light on the A-theory/B-theory debate in the philosophy of time by developing a novel answer to the question of whether one may reasonably claim that certain tensed sentences have irreducibly 'tensed' truth-conditions.

19.2.2 Two Modes of Temporal Individuation

Like individuals, also times emerge in our conceptualizations in two ways. Intentionally individuated times are based on an agent's temporal experience. Let us go back to our earlier example of Alice and the tomato (Sect. 19.1.2). Let V be

[5] Note that indeed any attempt to 'eternalize' the sentence (1b) must temporally fix not only the time of Mrs. Brown's not being at home, but also the time of Mary's belief.

[6] For denotations of temporal indexicals and date expressions, see Sect. 19.3.

[7] Some philosophers consider sentence tokens to be the primary truth-bearers. Tokens are occurrents: they appear in a specific scenario in a specific place at a specific time and they are brought into existence by a specific agent. Let τ be a token of the sentence type (1a). Let w_0 be the uniquely determined world in which τ occurs. The token τ occurs in w_0 at a certain instant s_0, but since (1a) contains a modal expression, the truth-condition of τ involves something more: a time $\mathfrak{t}_0$ with $s_0 = \mathfrak{t}_0(w_0)$. Actually, the truth-condition of τ would be something like (2b). The condition for the truth of τ would thus be precisely the same as the condition for the truth of (the content expressed by) the sentence type (1a) in w_0 at $\mathfrak{t}_0$, but here the token τ itself uniquely determines the context $(w_0, \mathfrak{t}_0)$. Even if τ is locally generated—it is brought about in the specific world w_0—it refers beyond w_0, since $\mathfrak{t}_0$ is a cross-world entity. I take truth-bearers to be contents expressed by sentence types in specific contexts. I will not further discuss the token-reflexive account of truth-conditions of tensed sentences, apart from shortly mentioning this idea in Sect. 19.4.1.

the set of contexts (w, t, l) relevant for Alice's perceptual experience. Perceptual experience is temporally determined: it pertains by its nature to the very time at which it takes place. This means that there is an intentionally individuated time $\mathfrak{p}$ that Alice experiences as the present moment such that for all $(w, t, l) \in V$, we have $t = \mathfrak{p}(w)$.[8] The content of Alice's perceptual experience involves two world lines: the intentional object $\mathfrak{j}$ (the tomato) and the intentionally individuated time $\mathfrak{p}$. Beside perception, other obvious modalities in connection with which there occur intentionally individuated times are memory, expectation, and belief. One of the main claims of this article is that the intentional mode of individuation extends much further. I shall maintain that what some philosophers have considered as ineliminably 'tensed' contents, are in the last analysis contents relying on intentionally individuated times.

We need to postulate physically individuated (agent-independent) times as well. Their cross-world behavior is regulated by physical laws. Such laws tell us, for example, what the state of a physical system would presently be, had certain initial conditions at its earlier stage been different in a certain specific way from what they in fact were. In a physical setting, the possibility of speaking of one and the same time being realized both in the actual circumstances and in counterfactual situations requires resorting to physically individuated times. Agent-independent times occur also outside the realm of scientific theories, when we in our everyday temporal discourse make use of regularities of slowly moving middle-sized bodies of our quotidian environment. For instance, the number of orbits the Earth has made around the Sun since the occurrence of the event chosen as the zero point of our chronology, together with information about the relative positions of the Sun and the Earth, serves to determine a specific moment—regarding which we can make not only factual but also counterfactual claims. In what follows, I will concentrate on physically individuated times of everyday life. Cases in which advanced physical theories are needed (the theory of relativity, quantum mechanics) will involve their own complications.

Considering once more Alice and the tomato, suppose Alice's perception takes place at noon on January 1, 1920, in Rome. The relevant moment, call it $\mathfrak{t}_0$, is a physically individuated time. As such, $\mathfrak{t}_0$ exhibits regularity in its behavior over the set V: for all $(w, t, l) \in V$, the instant $\mathfrak{t}_0(w)$ constitutes in w the midday of the date of January 1, 1920, in Rome. On the other hand, it is perfectly possible and indeed almost certain that Alice does not know her temporal location with absolute exactitude: there are contexts $(w, t, l) \in V$ such that $\mathfrak{t}_0(w) \neq \mathfrak{p}(w)$. The more contexts of this kind there are in V, the less precise, physically speaking, Alice's temporal perception is. Such an error in correctly locating oneself concerns the whole modal margin V: it is compatible with the fact that in the actual world w_0 we have $\mathfrak{p}(w_0) = \mathfrak{t}_0(w_0)$.

[8] I assume that there are no contexts $(w, t, l), (w', t', l') \in V$ such that $w = w'$ but $t \neq t'$. This assumption could be given up if *sets* of instants were allowed as values of temporal world lines.

In what follows I turn attention to phenomena which can be studied in the context of a propositional temporal language. We will need to consider only temporal world lines. The systematic study of cases which involve interactions of times and individuals—like the example of Alice and the tomato—is left to another occasion.

19.3 Formal Semantics

I formulate a propositional temporal language whose semantics makes use of temporal world lines. I do not aim to provide a *linguistic* analysis of natural-language temporal constructions. I only wish to find a way of logically representing a large variety of contents of sentences involving temporal expressions. If my analysis is fruitful, the conceptual distinctions I introduce will be of use also for an analysis with a more linguistic emphasis.

19.3.1 Basic Definitions

Let us fix a set **prop** of *propositional atoms* and a set **nom** of *temporal nominals*. Atoms represent unanalyzed states of affairs. In a model an atom is associated with the set of instants at which it is true. Nominals behave syntactically in many respects like atoms. In hybrid logic (see, e.g., [2]) nominals are formulas that denote possible circumstances of evaluation (worlds, instants).

We distinguish two types of nominals: nominals of type 1 denote instants, whereas nominals of type 2 denote times. The set **nom** is correspondingly divided into two disjoint parts: $\mathbf{nom}_1$ and $\mathbf{nom}_2$. I take *temporal indexicals* to be nominals of type 2: a temporal indexical denotes either a time provided by the non-linguistic context (this is the case notably with 'now') or a time defined in terms of such contextually given times (e.g., 'yesterday' as denoting the day preceding the day to which the denotation of 'now' belongs). Generally, we must be able to speak of the referent of an indexical in several worlds. Construing semantic values of temporal indexicals as temporal world lines predicts that they are potentially ambiguous between several readings: their denotation may be either intentionally or physically individuated. In the former case, furthermore, there may be several agents to be considered (e.g., the speaker, the addressee, or an agent to whom the sentence ascribes an attitude). I wish the language I formulate to be rather general and flexibly applicable in different situations. For example, date expressions such as 'May 7, 1906' might in some cases be best construed as denoting times and in other cases as denoting instants, and by recognizing the two types of nominals I avoid blocking potentially relevant interpretational possibilities.

I suppose there to be available a set M of *modality markers*, and a set A of *agent markers*. I apply the convention of writing **a** for the agent denoted by the agent marker *a*. The syntax of *propositional tense logic*, **TL**, is generated by the grammar

$$\varphi \ ::= p \mid \mathrm{n} \mid \underline{\mathrm{n}} \mid \neg\varphi \mid (\varphi \vee \varphi) \mid [m]\varphi \mid \mathrm{F}_a\,\varphi \mid \mathrm{P}_a\,\varphi \mid \mathrm{N}_a\,\varphi \mid \text{ꟻ}\varphi \mid \text{ꟼ}\varphi \mid \text{И}\varphi \mid @_{\underline{\mathrm{n}}}\varphi,$$

where $p \in \mathbf{prop}$, $\mathrm{n} \in \mathbf{nom}_1$, $\underline{\mathrm{n}} \in \mathbf{nom}_2$, $m \in \mathrm{M}$ and $a \in \mathrm{A}$. For nominals of type 2 we use underlined letters, so as to distinguish them from nominals of type 1. We assume that among nominals of type 2 there are the expressions ‘$\underline{\text{now}}$’ and ‘$\underline{\text{now(a)}}$’ for all $a \in \mathrm{A}$. Intuitively, ‘$\underline{\text{now}}$’ stands for the physical now-point of the context of utterance, while ‘$\underline{\text{now(a)}}$’ stands for the intentionally individuated time that the agent **a** experiences as present in the context of utterance.

The symbols ꟻ, ꟼ, И, F_a, P_a and N_a (with $a \in \mathrm{A}$) are *tense operators*. The semantic values of the operators ꟻ, ꟼ and И will be physically individuated times, and those of the operators F_a, P_a and N_a times intentionally individuated by the agent **a**. Here ꟻ and the F_a are *future tense* operators, ꟼ and the P_a are *past tense* operators, and И and the N_a *present tense* operators.[9] The expressions $@_{\underline{\mathrm{n}}}$ are *satisfaction operators*. We do not allow satisfaction operators relativized by nominals of type 1, since in our semantics the evaluation always needs a time component; denotations of nominals of type 1 are mere instants and would not serve to provide a time. Modal evaluation could not proceed if merely an instant was given instead of an entire time. When convenient, we may use the symbol $\langle m\rangle$ as an abbreviation of the sequence of symbols $\neg[m]\neg$. The syntax allows any number of modality markers, but for the needs of this paper it will suffice to concentrate on just one type of modality in any one formula. Depending on how we interpret that modality, it often becomes more convenient to write for example $\Box$ (alethic modality) or B_a (doxastic modality relative to agent **a**) or K_a (epistemic modality relative to agent **a**) instead of $[m]$. The connectives $\wedge$ and $\rightarrow$ are definable from the connectives $\neg$ and $\vee$ in the usual manner.

Worlds have an inner temporal structure: a set of instants ordered by an earlier-later relation.

Definition 19.1 (Temporal world, instants). A *temporal world* is a pair $w = \langle T, <_w\rangle$, where T is a non-empty set and $<_w$ be an irreflexive linear order on T (i.e., an irreflexive, transitive and trichotomous binary relation on T). The set T is called the *temporal domain* of w, denoted $\mathrm{temp}(w)$. Elements of temporal domains are termed *instants*. If W is a set of temporal worlds, we write $T(W)$ for the set $\bigcup_{w\in W} \mathrm{temp}(w)$.

I suppose that the earlier-later relation is irreflexive and linear, though more general relations could of course be considered. If within one and the same world there could be instants incomparable in terms of the earlier-later relation, the realization of a time in a world would generally be a set of instants rather

[9]The letters ‘N’ and its horizontal reflection ‘И’ are *not* meant to be reminiscent of the indexical word ‘now’. Temporal indexicals are represented by certain nominals $\underline{\mathrm{n}}$ of type 2, which may occur in a formula either by themselves—as subformulas—or as subscripts in the expression ‘$@_{\underline{\mathrm{n}}}$’.

than a single instant. Given our understanding of cross-world identity, no instant can belong to temporal domains of distinct worlds. This motivates the following definition.

Definition 19.2 (World of an instant). Let W be a collection of temporal worlds. For every instant $t \in T(W)$ there is a unique world $w \in W$ such that $t \in \text{temp}(w)$. We denote the unique world to whose temporal domain the instant t belongs by $\mathrm{w}(t)$, and call it the *world of the instant t*.

A time is naturally represented by a partial function assigning to each world in which it is realized an instant in the temporal domain of that world. However, to facilitate subsequent formulations, I take times to be total functions, with a special value $\star$ assigned to those worlds in which the time in question is not realized in the intended sense.

Definition 19.3 (Times, realization). If W is a set of temporal worlds and $\star \notin T(W)$, a *time* on W is a total function $\mathfrak{t}$ which to each world $w \in W$ assigns an object $\mathfrak{t}(w) \in \text{temp}(w) \cup \{\star\}$. We write time$(W)$ for the set of times on W. The object $\star$ is not an instant. If $\mathfrak{t}$ is a time and $\mathfrak{t}(w) \neq \star$, the instant $\mathfrak{t}(w)$ is the *realization* of $\mathfrak{t}$ in w. If, again, $\mathfrak{t}(w) = \star$, we say that the time $\mathfrak{t}$ not realized in w.

We will make use of the model-theoretic notions of skeleton and valuation.

Definition 19.4 (Availability, temporal skeleton). Let W be a set of temporal worlds. For every $s \in T(W)$, let $\mathfrak{T}_s$ be a (possibly empty) subset of time(W). Times $\mathfrak{t} \in \mathfrak{T}_s$ are said to be *available* at instant s. A *temporal skeleton* is a structure $\langle W, \mathfrak{T}\rangle$ with $\mathfrak{T} = \{\mathfrak{T}_s : s \in T(W)\}$, satisfying the following two conditions:

- *Actualism:* for all times $\mathfrak{t} \in \bigcup \mathfrak{T}$ and instants $s \in T(W)$, if $\mathfrak{t}$ is available at s, then $\mathfrak{t}$ is realized in the world of s. That is, if $\mathfrak{t} \in \mathfrak{T}_s$ and $w = \mathrm{w}(s)$, then $\mathfrak{t}(w) \neq \star$.
- *Strong separation:* for all $s_1, s_2 \in T(W)$, $\mathfrak{t}_1 \in \mathfrak{T}_{s_1}$ and $\mathfrak{t}_2 \in \mathfrak{T}_{s_2}$ with $\mathfrak{t}_1 \neq \mathfrak{t}_2$, the times $\mathfrak{t}_1$ and $\mathfrak{t}_2$ have no common realizations and are non-realized in the same worlds:

$$\mathfrak{t}_1(v) = \mathfrak{t}_2(v) \text{ iff } \mathfrak{t}_1(v) = \star \text{ or } \mathfrak{t}_2(v) = \star.$$

A skeleton $\langle W, \mathfrak{T}\rangle$ is *physical* (*intentional*) if all times in the sets $\mathfrak{T}_s$ with $s \in T(W)$ are physically (intentionally) individuated.

In our semantics the impact of the condition of actualism will be that those times $\mathfrak{t}$ over which we may quantify in w at s—whether they are physically or intentionally individuated—must be realized in w. The realization $\mathfrak{t}(w)$ of any such time available in w at s must, then, satisfy $\mathfrak{t}(w) < s$ or $\mathfrak{t}(w) = s$ or $s < \mathfrak{t}(w)$. Intentionally individuated *individuals* need not be realized when available (we may quantify over non-realized objects of belief or perceptual experience), but I take it that intentionally individuated *times* must nevertheless admit of a realization which bears a temporal relation to the instant at which we speak of those times. Intentionally individuated times are anchored to the world in which the agent resides. The requirement of strong separation has the consequence that for any given skeleton $\langle W, \mathfrak{T}\rangle$ and any $s \in T(W)$, the instant s is the realization of at most one time $\mathfrak{t} \in \mathfrak{T}$. This does not prevent there being two skeletons $\langle W_1, \mathfrak{T}_1\rangle$ and $\langle W_2, \mathfrak{T}_2\rangle$ with a common set of worlds $W_1 = W_2$, two worlds v and u, and two times $\mathfrak{t}_1 \in \mathfrak{T}_1$ and $\mathfrak{t}_2 \in \mathfrak{T}_2$ such that $\mathfrak{t}_1(v) = \mathfrak{t}_2(v)$ but $\mathfrak{t}_1(u) \neq \mathfrak{t}_2(u)$.

Definition 19.5 (Valuation). If W is a set of temporal worlds, a *valuation* on W is a function **Val** such that:

- If $p \in \mathbf{prop}$ and $w \in W$, then $\mathbf{Val}(p, w) \subseteq \mathrm{temp}(w)$;
- If $\mathrm{n} \in \mathbf{nom}_1$ and $w \in W$, then $\mathbf{Val}(\mathrm{n}, w) \in \mathrm{temp}(w)$;
- If $\underline{\mathrm{n}} \in \mathbf{nom}_2$, then $\mathbf{Val}(\underline{\mathrm{n}}) \in \mathrm{time}(W)$.

The denotation of a nominal of type 1 may change from world to world, while the denotation of a nominal of type 2 is invariant over worlds. In those cases in which we need nominals of type 2, their denotations are relative to the initial context of evaluation (the context of utterance), so we need not let the denotations vary depending on worlds possibly encountered during the evaluation.

Before being in a position to define the notion of temporal model, we need to distinguish variants of temporal skeletons.

19.3.2 Properties of Temporal Skeletons

Consider a skeleton $\langle W, \mathfrak{T}\rangle$. Let $w \in W$ and $s \in \mathrm{temp}(w)$. The notions of availability and realization must not be confused with each other. A time $\mathfrak{t}$ can be realized in w as the instant s without being available at s: we may have $\mathfrak{t}(w) = s$ without having $\mathfrak{t} \in \mathfrak{T}_s$. Conversely, $\mathfrak{t}$ can be available at s without being realized in w as the instant s: we can have $\mathfrak{t} \in \mathfrak{T}_s$ without having $\mathfrak{t}(w) = s$, though—because of the condition of actualism—$\mathfrak{t}$ must be realized somewhere in w.

Definition 19.6 (Transparent skeleton). A temporal skeleton $\langle W, \mathfrak{T}\rangle$ is *transparent* if realization implies availability in the following sense: for all $\mathfrak{t} \in \bigcup \mathfrak{T}$ and $s \in T(W)$,

$$\text{if } \mathfrak{t}(w) = s, \text{ then } \mathfrak{t} \in \mathfrak{T}_s.$$

A skeleton that is not transparent is *transcendent*: there are at least one time $\mathfrak{t}$, one world w and one instant $s \in \mathrm{temp}(w)$ such that the realization of $\mathfrak{t}$ in w is s, but $\mathfrak{t}$ is not available in s.

I take it to follow from the objective nature of physical times that physical skeletons are transparent. Intentional skeletons, again, may well be transcendent. For, suppose s_1 is an instant earlier than s_2 in world w. Suppose an agent **a** anticipates at s_1 that a certain event will take place: there is a time $\mathfrak{t} \in \mathfrak{T}^a_{s_1}$ such that $s_1 < \mathfrak{t}(w)$ and in all worlds v compatible with the agent's anticipation, the event takes place at $\mathfrak{t}(v)$. Now, it might happen that the event indeed will take place and that it takes place at s_2, so that we have $s_2 = \mathfrak{t}(w)$. However, by the time the instant s_2 becomes actual, the agent **a** may have totally forgotten what he or she was thinking at s_1; there is no reason why $\mathfrak{t}$ should belong to the set $\mathfrak{T}^a_{s_2}$. Independently of this issue, the realization of $\mathfrak{t}$ in w is s_2.

Definition 19.7 (Locally stable skeleton). A skeleton $\langle W, \mathfrak{T}\rangle$ is *locally stable*, if for all instants $t, s \in T(W)$ satisfying $\mathrm{w}(t) = \mathrm{w}(s)$, we have $\mathfrak{T}_t = \mathfrak{T}_s$.

In a locally stable skeleton, within each world the same times are available at all instants. However, the set of available times may vary with the world considered. Again, I take it to be a part of the objective character of physically individuated times that physical skeletons are locally stable. No similar assumption is made concerning intentional skeletons. Times available to an agent may well change from one instant to another even within one and the same world.

Realizations of two times may in principle be differently interrelated in distinct worlds.

Definition 19.8 (Respecting temporal order). A temporal skeleton $\langle W, \mathfrak{T}\rangle$ is said to *respect temporal order*, if for all instants $t, s \in T(W)$, all times $\mathfrak{t} \in \mathfrak{T}_t$ and $\mathfrak{s} \in \mathfrak{T}_s$ and all worlds $v, u \in W$ in which both $\mathfrak{t}$ and $\mathfrak{s}$ are realized, we have $\mathfrak{t}(v) <_v \mathfrak{s}(v)$ iff $\mathfrak{t}(u) <_u \mathfrak{s}(u)$.

I take physical skeletons to respect temporal order. Intentional skeletons need not do so. This reflects the fact that agents' temporal experience may be confused: for instance, an agent may recall that two events have taken place, and still he or she can be uncertain of their temporal order.

Definition 19.9 (Standard skeleton). A skeleton $\langle W, \mathfrak{T}\rangle$ is *standard* if it is transparent and locally stable, respects temporal order and satisfies: every $s \in T(W)$ is a realization of some $\mathfrak{t} \in \bigcup \mathfrak{T}$.

Because any standard skeleton $\langle W, \mathfrak{T}\rangle$ is transparent and every instant is a realization of a time, $\mathfrak{T}_s \neq \varnothing$ for all $s \in T(W)$. Since standard skeletons are locally stable, whenever $t, s \in T(W)$ and $\mathrm{w}(t) = \mathrm{w}(s)$, we have that $\mathfrak{T}_t = \mathfrak{T}_s$. In connection with standard skeletons we may, then, write $\mathfrak{T}_w$ for the unique set of times corresponding to all instants in the world $w = \mathrm{w}(t) = \mathrm{w}(s)$.

19.3.3 Semantics

Having the relevant definitions at our disposal, we fix the definition of temporal model and indicate how the language **TL** is evaluated over such models.

Definition 19.10 (Temporal model). A *temporal model* is a structure

$$\mathcal{M} = \langle W,\ \mathbf{time},\ (\mathfrak{T}_w)_{w\in W},\ (\mathfrak{T}^a_s)_{a\in \mathrm{A},\ s\in T(W)},\ (R^m_{\mathfrak{t}})_{m\in \mathrm{M},\ \mathfrak{t}\in \mathbf{time}},\ \mathbf{Val},\ \mathfrak{t}_0, \ldots, \mathfrak{t}_k\rangle,$$

where W is a set of temporal worlds, **time** is a set of times on W, each $\mathfrak{T}_w$ and each $\mathfrak{T}^a_s$ is a subset of **time**, each $R^m_\mathfrak{t}$ is a binary relation on W, **Val** is valuation on W, and each $\mathfrak{t}_i$ is an element of **time**. It is required that the components of a temporal model satisfy the following conditions:

- $\langle W, \mathfrak{T}\rangle$ is a standard physical skeleton.
- Each $\langle W, \mathfrak{T}^a\rangle$ is an intentional skeleton.
- The relations $R^m_\mathfrak{t}$ satisfy: if $uR^m_\mathfrak{t}u'$, then the time $\mathfrak{t}$ is realized in both u and u'.
- For every $\underline{\mathrm{n}} \in \mathbf{nom}_2$, $\mathbf{Val}(\underline{\mathrm{n}}) \in \{\mathfrak{t}_0, \dots, \mathfrak{t}_k\}$.
- $\mathbf{Val}(\underline{\mathrm{now}}) = \mathfrak{t}_0$ and $\mathfrak{t}_0$ is a physically individuated time.
- For each $a \in \mathrm{A}$, $\mathbf{Val}(\underline{\mathrm{now(a)}})$ is a time intentionally individuated by the agent **a**.

The $\mathfrak{t}_i$ are *time parameters* of the model, $\mathfrak{t}_0$ being its *first time parameter*. The sets W and **time** are respectively the *world domain* and the *time domain* of the model, denoted dom($\mathcal{M}$) and time($\mathcal{M}$).

We proceed to define recursively the satisfaction relation $\mathcal{M}, w, \mathfrak{t} \vDash \varphi$ for models $\mathcal{M}$, formulas $\varphi \in \mathbf{TL}$, worlds $w \in \mathrm{dom}(\mathcal{M})$ and times $\mathfrak{t} \in \mathrm{time}(\mathcal{M})$ realized in w.

$$
\begin{array}{rll}
\mathcal{M}, w, \mathfrak{t} \vDash p & \text{iff}: & \mathfrak{t}(w) \in \mathbf{Val}(p, w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \neg\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{t} \nvDash \psi \\
\mathcal{M}, w, \mathfrak{t} \vDash (\psi \vee \chi) & \text{iff}: & \mathcal{M}, w, \mathfrak{t} \vDash \psi \textit{ or } \mathcal{M}, w, \mathfrak{t} \vDash \chi \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{n} & \text{iff}: & \mathbf{Val}(\mathrm{n}, w) = \mathfrak{t}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \underline{\mathrm{n}} & \text{iff}: & \mathbf{Val}(\underline{\mathrm{n}})(w) = \mathfrak{t}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash [m]\psi & \text{iff}: & \text{for all } v \text{ with } wR^m_\mathfrak{t}v \text{ we have } \mathcal{M}, v, \mathfrak{t} \vDash \psi \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{F}\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{s} \vDash \psi \text{ for some } \mathfrak{s} \in \mathfrak{T}_w \text{ with } \mathfrak{t}(w) <_w \mathfrak{s}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{P}\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{s} \vDash \psi \text{ for some } \mathfrak{s} \in \mathfrak{T}_w \text{ with } \mathfrak{s}(w) <_w \mathfrak{t}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{N}\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{s} \vDash \psi \text{ for some } \mathfrak{s} \in \mathfrak{T}_w \text{ with } \mathfrak{t}(w) = \mathfrak{s}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{F}_a\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{s} \vDash \psi \text{ for some } \mathfrak{s} \in \mathfrak{T}^a_{\mathfrak{t}(w)} \text{ with } \mathfrak{t}(w) <_w \mathfrak{s}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{P}_a\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{s} \vDash \psi \text{ for some } \mathfrak{s} \in \mathfrak{T}^a_{\mathfrak{t}(w)} \text{ with } \mathfrak{s}(w) <_w \mathfrak{t}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash \mathrm{N}_a\psi & \text{iff}: & \mathcal{M}, w, \mathfrak{s} \vDash \psi \text{ for some } \mathfrak{s} \in \mathfrak{T}^a_{\mathfrak{t}(w)} \text{ with } \mathfrak{t}(w) = \mathfrak{s}(w) \\
\mathcal{M}, w, \mathfrak{t} \vDash @_{\underline{\mathrm{n}}}\psi & \text{iff}: & \mathbf{Val}(\underline{\mathrm{n}})(w) \neq \star \text{ and } \mathcal{M}, w, \mathbf{Val}(\underline{\mathrm{n}}) \vDash \psi.
\end{array}
$$

When the relation $\mathcal{M}, w, \mathfrak{t} \vDash \varphi$ prevails, we say that formula φ *holds* in model $\mathcal{M}$ in world w at time $\mathfrak{t}$. Once the satisfaction relation is defined, we explain in Definition 19.11 what is meant by the *truth* and *falsity* of a formula *in a model at a world*—as opposed to a formula holding or not holding in a model in a world at a time.

Let us say that an evaluation context $(\mathcal{M}, w, \mathfrak{t})$ has the *realization property*, if $\mathfrak{t}(w) \neq \star$. Now, if the evaluation of a formula ψ begins in a context with this property, the above clauses can never lead to a context not satisfying this property. First, the only clause which serves to shift the world parameter is the clause for modal operators $[m]$. The accessibility relations $R^m_\mathfrak{t}$ can only lead from a world in which $\mathfrak{t}$ is realized to a world in which $\mathfrak{t}$ is realized. Second, clauses for tense operators shift the time parameter, but they always introduce a time realized in the world of evaluation (the value $\star$ cannot be earlier, later or identical to a value distinct from $\star$). Third, the only remaining clause that shifts the time parameter is the clause for the satisfaction operators $@_{\underline{n}}$. Here the semantic clause is so formulated that the evaluation can only proceed if the denotation of $\underline{n}$ is realized in the world of evaluation.

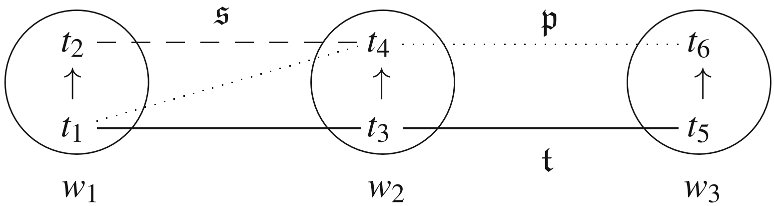
There are several details of the above semantics to which one should pay attention. First, the evaluation of atomic formulas is a purely local matter: whether p holds in w at $\mathfrak{t}$ depends only on how the valuation is defined on the world w at the instant $\mathfrak{t}(w)$. In particular, then, the satisfaction of atomic formulas does not depend on the times being realized, only on the instants realizing them. Second, the tense operators Ⅎ, ꟼ, И, F_a, P_a and N_a are semantically quantifiers that range over times $\mathfrak{s}$ bearing a certain specified relation to the time $\mathfrak{t}$ of evaluation. This relation is specified *locally*: in terms of how the instant realizing $\mathfrak{t}$ in w and the instant realizing $\mathfrak{s}$ in w are related according to the local temporal order $<_w$ of the world w of evaluation. Third, in w at $\mathfrak{t}$ the operator И ranges over the set $\{\mathfrak{s} \in \mathfrak{T}_w : \mathfrak{t}(w) = \mathfrak{s}(w)\}$. Since $\langle W, \mathfrak{T}\rangle$ is a standard skeleton, this set contains exactly one element. By contrast, in w at $\mathfrak{t}$ the range of the operator N_a is the set $\{\mathfrak{s} \in \mathfrak{T}^a_{\mathfrak{t}(w)} : \mathfrak{t}(w) = \mathfrak{s}(w)\}$. As the skeleton $\langle W, \mathfrak{T}^a\rangle$ need not be standard, this set may be empty—either because already $\mathfrak{T}^a_{\mathfrak{t}(w)}$ is empty, or because none of the times in $\mathfrak{T}^a_{\mathfrak{t}(w)}$ is realized as the instant $\mathfrak{t}(w)$. If the set $\{\mathfrak{s} \in \mathfrak{T}^a_{\mathfrak{t}(w)} : \mathfrak{t}(w) = \mathfrak{s}(w)\}$ is non-empty, it contains exactly one element, namely the time that the agent **a** experiences as present. Fourth, the tense operators Ⅎ, ꟼ, И, F_a, P_a and N_a are not syntactically quantifiers in the sense of being capable of binding occurrences of syntactically manifest variables. We could introduce a syntactic mechanism—the so-called $\downarrow$-binder as in hybrid logic [2]—which would compensate for this limitation by allowing to store times we encounter in the course of evaluation in a temporary 'memory' so as to be accessible later in the evaluation. Actually, I will make use of such a mechanism in an example to be considered in Sect. 19.5. Fifth, while in Priorean tense logic an atomic formula is by convention thought of as being present-tensed,[10] the same does not hold in my framework. Think of an atom p—as opposed to one of the formulas Иp or $\mathrm{N}_a p$—to which our attention may be turned either directly or through evaluating a complex formula. The atom p can be taken to be tenseless, or if one finds the idea of purely tenseless propositions to be an unwarranted abstraction, then one must say that p taken by itself is ambiguous, contextually disambiguated by the time of evaluation, which may be either physically or intentionally individuated. Sixth, if $\mathfrak{t}_0$ is the current physical time, $\mathrm{N}_a\psi$ holds in w at $\mathfrak{t}_0$ iff $\mathfrak{t}_0(w) = \mathfrak{p}(w)$ and ψ holds at $\mathfrak{p}$. Formulating the semantics in this way does not render the agent's temporal experience is infallible. While the realizations of $\mathfrak{p}$ and $\mathfrak{t}_0$ must coincide in w, the experienced time $\mathfrak{p}$ may differ to whatever extent from $\mathfrak{t}_0$ on worlds other than w.

The double character of nominals as formulas and denoting expressions deserves comments. In hybrid logic, nominals are considered as a limiting case of propo-

[10] Cf., e.g., [40, pp. 8–10], [41, pp. 14–15].

sitional atoms. They systematically stand for singleton sets of worlds, whereas propositional atoms may stand for any sets of worlds. Because they stand for one-element sets, they function as names; at the same time they are formulas. I treat temporal nominals similarly, though the more complex nature of my semantic framework brings in distinctions absent in usual formulations of hybrid logic. If we restrict attention not only to a fixed model $\mathcal{M}$ but also to a fixed world w, the semantic values of propositional atoms and nominals of both types will be sets of instants. First, an atom p stands for the set of instants $\mathbf{Val}(p, w) \subseteq \text{temp}(w)$. Second, a nominal n of type 1 stands for the singleton set $\{s\}$ with $s = \mathbf{Val}(\mathrm{n}, w)$. Third, if $\underline{\mathrm{n}}$ is a nominal of type 2 and $\mathfrak{t} = \mathbf{Val}(\underline{\mathrm{n}})$, then $\underline{\mathrm{n}}$ stands for the set of instants realizing the time $\mathfrak{t}$. At most one of these instants is an element of temp(w). Thus understood, the semantic values of propositional atoms are 'vertical': instants along the local earlier-later relation. By contrast, the semantic values of nominals of type 2 are 'horizontal': any two instants belong to distinct worlds. Using nominals we can make claims about the (local) identity of the instant $\mathfrak{t}(w)$ of evaluation. A nominal n of type 1 holds in w at $\mathfrak{t}$ iff the instant $\mathfrak{t}(w)$ equals the denotation $\mathbf{Val}(\mathrm{n}, w)$ of n in w. If $\underline{\mathrm{n}}$ is a nominal of type 2, its denotation $\mathbf{Val}(\underline{\mathrm{n}})$ is a temporal world line; this is why in the clause for $\underline{\mathrm{n}}$ we apply the time $\mathbf{Val}(\underline{\mathrm{n}})$ to the world w so as to obtain an instant $\mathbf{Val}(\underline{\mathrm{n}})(w)$. The nominal $\underline{\mathrm{n}}$ holds in w at $\mathfrak{t}$ iff the instant $\mathfrak{t}(w)$ equals the instant $\mathbf{Val}(\underline{\mathrm{n}})(w)$.

Let us illustrate the semantics by examples. Let $W = \{w_1, w_2, w_3\}$ with $w_1 \neq w_2 \neq w_3 \neq w_1$ and $\text{temp}(w_1) = \{t_1, t_2\}$, $\text{temp}(w_2) = \{t_3, t_4\}$ and $\text{temp}(w_3) = \{t_5, t_6\}$, each temporal domain consisting of two distinct instants. Let $<_{w_1} = \{(t_1, t_2)\}$, $<_{w_2} = \{(t_3, t_4)\}$, and $<_{w_3} = \{(t_5, t_6)\}$. Let $\mathfrak{T}_{w_1} = \mathfrak{T}_{w_2} = \{\mathfrak{t}, \mathfrak{s}\}$ and $\mathfrak{T}_{w_3} = \{\mathfrak{t}\}$, where the physical times $\mathfrak{t}$ and $\mathfrak{s}$ satisfy: $\mathfrak{t}(w_1) = t_1$ and $\mathfrak{t}(w_2) = t_3$ and $\mathfrak{t}(w_3) = t_5$, and $\mathfrak{s}(w_1) = t_2$ and $\mathfrak{s}(w_2) = t_4$ and $\mathfrak{s}(w_3) = \star$. Further, let $\mathfrak{T}^a_{t_1} = \mathfrak{T}^a_{t_4} = \mathfrak{T}^a_{t_6} = \{\mathfrak{p}\}$, where the intentionally individuated time $\mathfrak{p}$ satisfies $\mathfrak{p}(w_1) = t_1$ and $\mathfrak{p}(w_2) = t_4$ and $\mathfrak{p}(w_3) = t_6$.



Let $R_{\mathfrak{t}} = \{(w_1, w_2), (w_1, w_3)\} = R_{\mathfrak{p}}$ and $R_{\mathfrak{s}} = \{(w_1, w_2)\}$ be the accessibility relations indicating the worlds compatible with the belief of the agent **a** relative to the three times. Finally, let $\mathbf{Val}(\underline{\mathrm{now}}) = \mathfrak{t}$, $\mathbf{Val}(r, w_1) = \{t_1\}$, $\mathbf{Val}(p, w_2) = \{t_3\} = \mathbf{Val}(q, w_2)$, $\mathbf{Val}(q, w_3) = \{t_5\}$, and $\mathbf{Val}(p, w_1) = \mathbf{Val}(p, w_3) = \mathbf{Val}(q, w_1) = \mathbf{Val}(r, w_2) = \mathbf{Val}(r, w_3) = \varnothing$. Write $\mathcal{M} = \langle W,\ \{\mathfrak{t}, \mathfrak{s}, \mathfrak{p}\},\ (\mathfrak{T}_w)_{w \in W},\ (\mathfrak{T}^a_s)_{s \in T(W)},\ \{R_{\mathfrak{t}}, R_{\mathfrak{s}}, R_{\mathfrak{p}}\},\ \mathbf{Val},\ \mathfrak{t}\rangle$.

(a) We have $\mathcal{M}, w_1, \mathfrak{t} \models V\!\!\!\mid B_a q$. For, there is a physically individuated time $\mathfrak{h} \in \mathfrak{T}_{w_1}$ with $\mathfrak{h}(w_1) = t_1 = \mathfrak{t}(w_1)$ such that for all worlds v accessible along $R_{\mathfrak{h}}$ from

w_1, we have: $\mathcal{M}, v, \mathfrak{h} \vDash q$. Actually, $\mathfrak{t}$ itself is such a time $\mathfrak{h}$. Namely, the worlds accessible along $R_\mathfrak{t}$ from w_1 are w_2 and w_3, and indeed $\mathfrak{t}(w_2) = t_3 \in \mathbf{Val}(q, w_2)$ and $\mathfrak{t}(w_3) = t_5 \in \mathbf{Val}(q, w_3)$. This means that presently **a** believes that q, where 'presently' is understood in the objective physical sense.

(b) We do *not* have $\mathcal{M}, w_1, \mathfrak{t} \vDash \mathrm{N}_a B_a q$. Namely, there is only one time intentionally individuated by **a** which agrees with $\mathfrak{t}$ on w_1, namely $\mathfrak{p}$ satisfying indeed $\mathfrak{p}(w_1) = t_1 = \mathfrak{t}(w_1)$, but $\mathcal{M}, w_1, \mathfrak{p} \nvDash B_a q$. This follows from the fact that $w_1 R_\mathfrak{p} w_2$ but $\mathcal{M}, w_2, \mathfrak{p} \nvDash q$, as $\mathfrak{p}(w_2) = t_4 \notin \mathbf{Val}(q, w_2)$. What this means is that presently **a** does not believe that q, where 'presently' is understood in the internal sense of what the agent takes to be the present time. This is possible even though $\text{И} B_a q$ holds in w_1 at $\mathfrak{t}$: the agent mistakes the physical time $\mathfrak{t}$ for the time $\mathfrak{p}$.

(c) Write ⅁ for the dual of Ⅎ, so that $\text{⅁}\psi$ means $\neg\text{Ⅎ}\neg\psi$: 'always in the future ψ'. We have: $\mathcal{M}, w_1, \mathfrak{t} \vDash (\text{⅁}\neg r \wedge \text{⅁}@_{\underline{\text{now}}} r)$. First, $\mathcal{M}, w_1, \mathfrak{t} \vDash \text{⅁}\neg r$ holds, since for every physical time $\mathfrak{h}$ satisfying $\mathfrak{t}(w_1) <_{w_1} \mathfrak{h}(w_1)$ we have that r fails in w_1 at $\mathfrak{h}(w_1)$. There is just one such time, namely $\mathfrak{s}$, and $\mathfrak{s}(w_1) = t_2 \notin \mathbf{Val}(r, w_1)$. Second, because we have $\mathfrak{t}(w_1) = t_1 \in \mathbf{Val}(r, w_1)$ and therefore $\mathcal{M}, w_1, \mathfrak{t} \vDash r$, we also have $\mathcal{M}, w_1, \mathfrak{t} \vDash \text{⅁}@_{\underline{\text{now}}} r$. Since r holds at $\mathfrak{t} = \mathbf{Val}(\underline{\text{now}})$, it will always be the case that r holds at $\mathfrak{t}$.

(d) On the one hand, $B_a @_{\underline{\text{now}}} p$ does not hold in w_1 at $\mathfrak{t}$. For, $w_1 R_\mathfrak{t} w_3$ and $\mathcal{M}, w_3, \mathfrak{t} \nvDash p$, since $\mathfrak{t}(w_3) = t_5 \notin \mathbf{Val}(p, w_3)$. On the other hand, $\mathcal{M}, w_1, \mathfrak{t} \vDash \text{Ⅎ} B_a @_{\underline{\text{now}}} p$. Namely, there is a physical time $\mathfrak{h}$ satisfying $\mathfrak{t}(w_1) <_{w_1} \mathfrak{h}(w_1)$ such that $\mathcal{M}, w_1, \mathfrak{h} \vDash B_a @_{\underline{\text{now}}} p$. In fact, $\mathfrak{s}$ is such a physical time: for all worlds v with $w_1 R_\mathfrak{s} v$, we have $\mathcal{M}, v, \mathfrak{s} \vDash @_{\underline{\text{now}}} p$. Actually, the only such world v is w_2, and indeed $\mathcal{M}, w_2, \mathfrak{s} \vDash @_{\underline{\text{now}}} p$ because $\mathbf{Val}(\underline{\text{now}})(w_2) = \mathfrak{t}(w_2) = t_3 \in \mathbf{Val}(p, w_2)$. Even if an agent does not believe that p holds at $\mathfrak{t}$, the agent may come to believe that it does.

Here is another example. Write $\rotatebox[origin=c]{90}{H}$ for the dual of ꟼ, so that $\rotatebox[origin=c]{90}{H}\psi$ means $\neg\text{ꟼ}\neg\psi$: 'always in the past ψ'. Similarly, write H_a for the dual of P_a. Let $\mathcal{M}$ be a model and $\mathfrak{t}_0$ its first time parameter. Suppose $\mathcal{M}, w, \mathfrak{t}_0 \vDash (\text{И}\Box p \wedge \rotatebox[origin=c]{90}{H}\Box\text{⅁}\neg p)$. It follows that for all *physically* individuated times $\mathfrak{s}$ with $\mathfrak{s}(w) <_w \mathfrak{t}_0(w)$, the sets of worlds $\{v : R_{\mathfrak{t}_0}(w, v)\}$ and $\{v : R_\mathfrak{s}(w, v)\}$ are disjoint. Namely, suppose towards a contradiction that there is a physical time $\mathfrak{s}$ and a world u such that $\mathfrak{s}(w) <_w \mathfrak{t}_0(w)$ and we have both $R_{\mathfrak{t}_0}(w, u)$ and $R_\mathfrak{s}(w, u)$. Then, on the one hand we have $\mathcal{M}, u, \mathfrak{t}_0 \vDash p$, and on the other hand $\mathcal{M}, u, \mathfrak{s} \vDash \text{⅁}\neg p$. However, since the physical skeleton of $\mathcal{M}$ respects temporal order, we have $\mathfrak{s}(u) <_u \mathfrak{t}_0(u)$. Because this skeleton is transparent, the time $\mathfrak{t}_0$ belongs to $\mathfrak{T}_u$ (instead of merely being realized in u). It follows that $\mathcal{M}, u, \mathfrak{t}_0 \vDash \neg p$. This is impossible.

Let us move on to consider the formula $(\mathrm{N}_a\Box p \wedge \mathrm{H}_a\Box\mathrm{G}_a\neg p)$. It is compatible with this formula's holding in $\mathcal{M}$ in w at $\mathfrak{t}_0$ that there are times $\mathfrak{s}, \mathfrak{t} \in \mathfrak{T}^a_{\mathfrak{t}_0(w)}$ *intentionally* individuated by the agent **a** such that $\mathfrak{s}(w) <_w \mathfrak{t}(w) = \mathfrak{t}_0(w)$ and the sets of worlds $\{v : R_\mathfrak{t}(w, v)\}$ and $\{v : R_\mathfrak{s}(w, v)\}$ are *not* disjoint, i.e., there is a world u such that $R_\mathfrak{t}(w, u)$ and $R_\mathfrak{s}(w, u)$. In such a model we will have $\mathcal{M}, w, \mathfrak{t} \vDash \Box p$ and $\mathcal{M}, w, \mathfrak{s} \vDash \Box\mathrm{G}_a\neg p$. Therefore the world u must satisfy $\mathcal{M}, u, \mathfrak{t} \vDash p$ and

$\mathcal{M}, u, \mathfrak{s} \vDash G_a \neg p$. That is, $\mathfrak{t}(u) \in \mathbf{Val}(p, u)$ but for all $\mathfrak{r} \in \mathfrak{T}^a_{\mathfrak{s}(u)}$ with $\mathfrak{s}(u) <_u \mathfrak{r}(u)$ we have $\mathfrak{r}(u) \notin \mathbf{Val}(p, u)$. This would be a contradiction if the time $\mathfrak{t}$ belonged to the set $\mathfrak{T}^a_{\mathfrak{s}(u)}$ *and* satisfied the condition $\mathfrak{s}(u) <_u \mathfrak{t}(u)$. However, neither of these conditions needs to hold. First, $\mathfrak{t}$ need not belong to the set $\mathfrak{T}^a_{\mathfrak{s}(u)}$: since $\mathfrak{t}$ is intentionally individuated, its being realized in u does not imply that there is an instant in temp(u) at which $\mathfrak{t}$ is available. In particular, $\mathfrak{s}(u)$ need not be such an instant. Second, even if $\mathfrak{t}$ were available at $\mathfrak{s}(u)$, we could have $\mathfrak{t}(u) <_u \mathfrak{s}(u)$ while having $\mathfrak{s}(w) <_w \mathfrak{t}(w)$, since intentional skeletons need not respect temporal order.

19.3.4 Truth and Falsity

Using the satisfaction relation $\mathcal{M}, w, \mathfrak{t} \vDash \varphi$, we define the basic semantic relation $\mathcal{M}, w \vDash \varphi$. When this relation holds, we say that formula φ is true in model $\mathcal{M}$ at world w. While the relation $\mathcal{M}, w \vDash \varphi$ is not *explicitly* relative to a time, it should be thought of as being relative to the first time parameter $\mathfrak{t}_0$ of the model. This temporal parameter plays the role of the physical now-point of the non-linguistic context represented by $\mathcal{M}$.[11]

Definition 19.11 (Truth, falsity). Let φ be a formula of **TL**. Let $\mathcal{M}$ be a temporal model, with $w \in \text{dom}(\mathcal{M})$. By definition $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}, w, \mathfrak{t}_0 \models \varphi$. When this condition holds, we say that φ is *true* in $\mathcal{M}$ at w. If $\mathcal{M}, w \not\models \varphi$, we say φ is *false* in $\mathcal{M}$ at w. Derivatively, if $\mathcal{M}, w, \mathfrak{t} \models \varphi$, we say that φ is *true* in $\mathcal{M}$ in w at $\mathfrak{t}$, and if $\mathcal{M}, w, \mathfrak{t} \not\models \varphi$, we say φ is *false* in $\mathcal{M}$ in w at $\mathfrak{t}$.

The rationale of the above definition is that we wish the basic semantic relation to be general enough to cover cases where we need a 'time of utterance'—the current physical time of the non-linguistic context—in order to be able to phrase the truth-condition of a formula. Such a time must be fixed at the outset, before the evaluation of a formula begins.

It is useful to note the following equivalences concerning the notions of truth at a world and truth in a world at a time. Let $\mathcal{M}$ be a model, with $\mathfrak{t}_0$ as its first time parameter. For all formulas $\varphi \in$ **TL**, worlds $w \in \text{dom}(\mathcal{M})$ and times $\mathfrak{t} \in \text{time}(\mathcal{M})$ with $\mathfrak{t}(w) = \mathfrak{t}_0(w)$, we have:

$$\mathcal{M}, w \vDash \varphi \text{ iff } \mathcal{M}, w \vDash \not{N}\varphi \text{ iff } \mathcal{M}, w, \mathfrak{t} \vDash \not{N}\varphi.$$

[11] Discussing the indexical 'now', Kamp [19] formulates a semantics using double indexing: one temporal parameter remains fixed and keeps track of the initial time of evaluation (the now-point), while the other temporal parameter changes in accordance with the semantics of the different temporal constructions evaluated. Our parameter $\mathfrak{t}_0$ provided by the model will allow us to capture some phenomena related to indexicality, and it plays in effect the role of the extra time parameter that Kamp postulates for enabling an analysis of temporal indexicality.

First, $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}, w, \mathfrak{t}_0 \models \varphi$ iff $\mathcal{M}, w, \mathfrak{t}_0 \models \not{N}\varphi$ iff $\mathcal{M}, w \models \not{N}\varphi$. The first and the third equivalence hold by the definition of truth in a model at a world, and the second equivalence by the semantics of the operator $\not{N}$ and the fact that $\mathfrak{t}_0$ is the only physical time $\mathfrak{s}$ available in w and satisfying $\mathfrak{s}(w) = \mathfrak{t}_0(w)$. Second, let $\mathfrak{t}$ be an arbitrary time satisfying $\mathfrak{t}(w) = \mathfrak{t}_0(w)$; we suppose nothing of its mode of individuation. We have $\mathcal{M}, w, \mathfrak{t} \models \not{N}\varphi$ iff there is a physically individuated time $\mathfrak{s}$ available in w and coinciding with $\mathfrak{t}$ on w such that $\mathcal{M}, w, \mathfrak{s} \models \varphi$. By the strong separation of the physical skeleton of $\mathcal{M}$, the only such time $\mathfrak{s}$ is $\mathfrak{t}_0$. We have already observed that $\mathcal{M}, w, \mathfrak{t}_0 \models \varphi$ iff $\mathcal{M}, w \models \not{N}\varphi$, so we may infer that $\mathcal{M}, w \models \not{N}\varphi$ iff $\mathcal{M}, w, \mathfrak{t} \models \not{N}\varphi$.

Despite what just seen, the formulas φ and $\not{N}\varphi$ are not interchangeable salva veritate in arbitrary syntactic positions within a larger formula. For example, while $\Diamond p$ and $\not{N}\Diamond p$ are equivalent, $\mathrm{N}_a\Diamond p$ and $\mathrm{N}_a\not{N}\Diamond p$ are not.[12] Let $\mathcal{M}$ be a model whose first time parameter is $\mathfrak{t}_0$ and whose domain consists of w and v, with $\mathrm{temp}(w) = \{t_0\}$ and $\mathrm{temp}(v) = \{t, s\}$, where $t \neq s$. Let $\mathfrak{T}_w = \mathfrak{T}_v = \{\mathfrak{t}_0\}$ and $\mathfrak{T}^a_t = \{\mathfrak{s}\}$ and $\mathfrak{T}^a_s = \emptyset$. Let $\mathfrak{t}_0(w) = t = \mathfrak{s}(w)$, and let $\mathfrak{t}_0(v) = t$ while $\mathfrak{s}(v) = s$. Let $R_{\mathfrak{t}_0} = \{(w, v)\} = R_{\mathfrak{s}}$. Finally, let $\mathbf{Val}(p, v) = \{s\}$. Now, $\mathcal{M}, w, \mathfrak{s} \models \Diamond p$, because $wR_{\mathfrak{s}}v$ and $\mathfrak{s}(v) = s \in \mathbf{Val}(p, v)$. Because $\mathfrak{s}$ is a time intentionally individuated by the agent **a** and available in w at $t = \mathfrak{t}_0(w) = \mathfrak{s}(w)$, it follows that $\mathcal{M}, w, \mathfrak{t}_0 \models \mathrm{N}_a\Diamond p$, that is, $\mathcal{M}, w \models \mathrm{N}_a\Diamond p$. On the other hand, $\mathcal{M}, w \not\models \mathrm{N}_a\not{N}\Diamond p$. For, otherwise we would have $\mathcal{M}, w, \mathfrak{s} \models \not{N}\Diamond p$ and therefore $\mathcal{M}, w, \mathfrak{t}_0 \models \Diamond p$ and further $\mathcal{M}, v, \mathfrak{t}_0 \models p$, which is not the case, as $\mathfrak{t}_0(v) = t \notin \mathbf{Val}(p, v)$.

Let us see in which way our logical formalism can be useful in connection with natural language semantics.[13] My framework predicts that the modifier 'presently' is potentially ambiguous. It can either be interpreted in terms of the current physical time ($\not{N}$) or else in terms of the time that a contextually relevant agent **a** experiences as present (N_a). Further, since my semantic theory analyzes temporal indexicals modally, as denoting world lines, there is a potential ambiguity with indexicals such as 'now': their semantic values can be either physically or intentionally individuated temporal world lines. For disambiguation, we use the syntactic distinction between '$\underline{\text{now}}$' and '$\underline{\text{now(a)}}$' with $a \in \mathrm{A}$. The times over which the operators $\not{N}$ and N_a range depend on the world of evaluation. In particular, $\not{N}$ and $@_{\underline{\text{now}}}$ do not in general have the same semantic effect: when evaluated in w at $\mathfrak{t}$, the latter systematically switches the time of evaluation to the first time parameter ($\mathfrak{t}_0$) of the model, while the former leads us to consider the (unique) physically individuated time that coincides with $\mathfrak{t}$ on w; this time need not be the same as $\mathfrak{t}_0$. For a simple example, consider

[12] By saying that φ and ψ are *equivalent*, we mean that they satisfy $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}, w \models \psi$ for all models $\mathcal{M}$ and worlds $w \in \mathrm{dom}(\mathcal{M})$.

[13] In future work we must address the systematic question of whether and how a *compositional* natural language semantics can be developed on the basis of the ideas presented in this paper. The details are not entirely obvious. For example, under one natural reading, 'It rains' has the form $\not{N}p$, while the corresponding reading of 'Mary believes that it rains' has the form $\not{N}B_{\mathrm{Mary}}p$, not $B_{\mathrm{Mary}}\not{N}p$. Special attention must be paid to the interactions of representations of grammatical tenses and expressions for propositional attitudes.

the formulas $\text{ꟼ}\text{И}p$ and $\text{ꟼ}@_{\underline{\text{now}}}p$. The former is equivalent to $\text{ꟼ}p$ and the latter to $(\text{ꟼ}\top \wedge p)$.[14] When evaluating $\text{ꟼ}\text{И}p$ in $\mathcal{M}$ at w, the semantic value of the operator И equals the semantic value $\mathfrak{s}$ of ꟼ—which is a certain physically individuated time satisfying $\mathfrak{s}(w) <_w \mathfrak{t}_0(w)$. On the other hand, when evaluating the formula $\text{ꟼ}@_{\underline{\text{now}}}p$ in $\mathcal{M}$ at w, the operator $@_{\underline{\text{now}}}$ switches the evaluation time to the value $\mathfrak{t}_0$ of '$\underline{\text{now}}$' which cannot be equal to $\mathfrak{s}$.

My semantics predicts that the English sentence

4. Mary believes that it rains

is ambiguous between various possible readings. In order to specify the readings, I indicate formulas whose evaluation is implicitly relative to the first time parameter $\mathfrak{t}_0$ of the model of evaluation. Suppose it is John who is ascribing to Mary the attitude expressed by our sentence. At least for each of the following logical forms there is a possible reading of this sentence in this situation:

5. (a) $\text{И}B_{\text{Mary}}\,p$ (or, equivalently, $B_{\text{Mary}}\,p$)
   (b) $\mathrm{N}_{\text{Mary}}B_{\text{Mary}}\,p$
   (c) $\mathrm{N}_{\text{John}}B_{\text{Mary}}\,p$.

I will refer to (5a) as a *B-theoretic reading*.[15] It ascribes to Mary a belief which is temporally structured according to the actual physical time $\mathfrak{t}_0$: in every world $v \in V$, it rains in v at $\mathfrak{t}_0(v)$, where V is the set of worlds compatible with Mary's beliefs in the actual world at $\mathfrak{t}_0$. Under which condition is Mary's belief thus structured, then? Normally this would require that she has an accurate opinion of the actual physical time $\mathfrak{t}_0$ over the modal margin that corresponds to the worlds compatible with her belief, in the sense that for all worlds v compatible with all that Mary believes in w_0 at $\mathfrak{t}_0$, we have $\mathfrak{t}_0(v) = \mathfrak{p}(v)$, where $\mathfrak{p}$ is an intentionally individuated time satisfying $\mathfrak{p} \in \mathfrak{T}^{\text{Mary}}_{\mathfrak{t}_0(w_0)}$ and $\mathfrak{t}_0(w_0) = \mathfrak{p}(w_0)$. Since intentional skeletons are strongly separated, if there is such a time, there is exactly one such time. In principle this ascription could be correct even if there was no such time $\mathfrak{p}$, i.e., if the set $\{\mathfrak{p} \in \mathfrak{T}^{\text{Mary}}_{\mathfrak{t}_0(w_0)} : \mathfrak{t}_0(w_0) = \mathfrak{p}(w_0)\}$ was empty. Mary might have a belief which just happens to be structured via the current physical time $\mathfrak{t}_0$, although no time intentionally individuated by Mary coincides with $\mathfrak{t}_0$ over the relevant margin. This is a rather remote possibility, but can be discerned as a separate alternative in my framework.

The reading (5b) provides an 'internal' description of what Mary believes. It attributes to Mary a belief which is temporally structured according to the intentionally individuated time $\mathfrak{p}$ that Mary experiences as present: in every world v compatible with Mary's belief in the actual world at $\mathfrak{t}_0$, it rains in v at $\mathfrak{p}(v)$. It should be observed that John could in fact correctly attribute to Mary what is expressed by (5b) even if he had no clue as to what Mary takes to be the present time. Of

[14] We let $\top$ be an abbreviation of $(q \vee \neg q)$, where q is a fixed atom. In addition to saying that p is currently true, $\text{ꟼ}@_{\underline{\text{now}}}p$ says that at least one physically individuated time is realized in the past.

[15] For A- and B-theories, see Sect. 19.4.

course in order for such an attribution to be *justified*, John needs some such clue. But surely the attribution might be correct while unjustified. Construed as (5c), the sentence offers likewise an 'internal' description, but in this case John's temporal experience is transferred to Mary in the sense that (5c) attributes to Mary a belief which is temporally structured according to what *John* takes to be the present time: if $\mathfrak{s} \in \mathfrak{T}^{\text{John}}_{\mathfrak{t}_0(w_0)}$ and $\mathfrak{s}(w_0) = \mathfrak{t}_0(w_0)$, then in every world v compatible with Mary's belief in the actual world at $\mathfrak{t}_0$, it rains in v at $\mathfrak{s}(v)$.

I will say that (5b) and (5c) ascribe to Mary a *tensed belief* and that these are *A-theoretic readings* of the English sentence 'Mary believes that it rains'. An A-theoretic reading ascribes to Mary a belief structured by an intentionally individuated time. It goes without saying that it is not entirely easy to ascertain whether a condition such as those expressed by the readings (5a–c) indeed holds.[16] They are not simple extensional claims. They are not even simple modal claims: it is not enough to consider a number of alternative worlds one by one and reflect on whether the formula p is true in each of them separately. Yet one could, supposing that Mary is sufficiently collaborative, come to know whether for example (5a) holds. This would require finding out whether the time Mary experiences as present differs from the relevant physical time. The more Mary's beliefs leave room for alternatives, the more likely it is that these temporal world lines differ from one another.

19.4 The A-Theory Analyzed

19.4.1 The A-Theory and the B-Theory

In the philosophy of time, there is a persistent debate between those who hold that there are ineliminable properties such as *past*, *present* and *future* and those who deny the reality of these properties. The respective views are called the *A-theory of time* and the *B-theory of time* (or the *tensed* and the *tenseless* theory), with reference to a distinction McTaggart [26] introduced when speaking of two ways of viewing interrelations of positions in time. Among influential A-theorists are William Lane Craig and Quentin Smith [7, 55], while Hugh Mellor and L. Nathan Oaklander [27, 31, 32] are well-known B-theorists.[17] The A-theory/B-theory debate has repercussions on the semantics of temporal expressions as well as on issues of temporal ontology.

[16] Then again, most counterfactual claims that we employ in our everyday lives are impossible to verify (e.g., 'Had *X* occurred, *Y* would have performed action *Z*'), and yet they play an important role in our practical decision making.

[17] Oaklander has recently given up the B-theory; now he defends what he calls the R-theory of time (the Russellian theory of time), see [35]. Due to Smith's criticism [55], Mellor has changed the view he presented in [27]—see [28]—but he continues to defend a variant of the B-theory.

We may follow Broad [3] and Geach [11] in distinguishing two types of characteristics. *Being past*, *being present*, *being future*, *being yesterday* and *being ten years ago* are A-characteristics: they are qualities that can only be ascribed to entities from a fixed point of view which is taken to be the present. Sometimes the A-theorist speaks of 'A-determinations' when referring to A-characteristics. *Being earlier than*, *being later than*, *lasting an hour* and *being ten years apart in birthdays*, again, are B-characteristics: possessing such characteristics is not relative to a particular now-point. A-characteristics are unary relations (properties) and they are perspectival by nature. B-characteristics are binary relations and they are context-independent. Whereas the A-theory maintains that A-characteristics are objective and independent features of reality, according to the B-theory tenseless B-characteristics are sufficient to account for tensed language and timely action. The B-theory has different variants.

The so-called *Old B-theory* claimed that tensed sentences can be translated without loss of meaning into tenseless sentences, either by replacing temporal constructions by an explicit specification of dates (the date theory advocated by Russell [48, 49], Frege [10] and Quine [43]), or else by assuming that primary truth-bearers are sentence tokens and replacing in a token τ expressions for A-characteristics by expressions for B-characteristics relativized to the occurrence time of τ (the token-reflexive theory advocated by Reichenbach [46] and Smart [52, 53]). Whereas according to the A-theory the sentence 'Mrs. Brown is not at home', uttered on May 7, 1906, ascribes *presentness* to the event of Mrs. Brown's not being at home, the date version of the Old B-theory holds that this sentence is translated by the atemporal sentence 'Mrs. Brown *is* not at home on May 7, 1906' and the token-reflexive version says that any token τ of this sentence is translated by the atemporal sentence 'Mrs. Brown *is* not at home at the time of utterance of τ'.[18] The putative translations thus obtained appear to be wrong (cf., e.g., Perry [38]): upon hearing someone utter 'Mrs. Brown is not at home', one gains information not gained by hearing an utterance of 'Mrs. Brown is not at home on May 7, 1906' (and vice versa). Also, most sentence types express contents whose truth does not require that these sentence types have tokens. The truth of the content expressed by a sentence type such as 'There are no tokens' even precludes the existence of true tokens.[19] For reasons such as these, both variants of the Old B-theory are widely considered misguided.

---

[18]Recall from Sect. 19.2.1 the convention of using italics to mark a tenseless verb form. Philosophers engaged in the A-theory/B-theory debate tend to phrase their discussion in terms of *events*. It is not always clear in this connection what sorts of entities events are supposed to be and how they are related to propositional contents. Conceivably this way of speaking is based on a hasty generalization from cases in which a predicate is attributed to a subject. Or, is there—over and above the content *that* $\forall x\ \exists y$ *such that x looks for y*—also an event of everyone's looking for someone. In the interest of space I do not comment on this issue further, and allow speaking of events when discussing the A-theory/B-theory distinction. I will reject both theories anyway.

[19]For the token-reflexive view on truth-conditions, cf. Footnote 7 in Sect. 19.2.1.

The so-called *New B-theory* gives up the translation thesis and admits the ineliminability of tensed language and thought, but claims that truth-conditions for tokens of tensed sentences can nevertheless be provided in tenseless terms and that therefore there is no need to postulate anything like ontological tenses (Dyke [8]; Mellor [27, 28]; Mozersky [29, 30]; Oaklander [31–34]; Paul [36]). According to this theory, the truth-condition of a token τ of the sentence 'Mrs. Brown is not at home' is the tenseless condition that the token τ is simultaneous with the relevant event of Mrs. Brown's not being at home. The New B-theory has been criticized in different ways. Smith [54] has argued that the theory leads to the absurd conclusion that simultaneous tokens of the same sentence differ in meaning (as they invoke distinct tokens). A defense has been put forward by Oaklander [32] and Mozersky [29]; in one way or another both authors resort to the New Theory of Reference with its idea of direct reference. The theory of direct reference is rejected on independent grounds in my framework: it operates with an unproblematized notion of identity, takes for granted that one and the same object can be a denizen of several worlds (which I take to be a mistake), and proceeds from the idea that fixing an element of the domain of the actual world as the referent of a term allows us to speak of that same thing as existing in other worlds as well; for a critique of this type, see [15, pp. 19–34], [16, 57]. Concerning defenders of the New B-theory, it must be noted, as Craig [7] does, that direct reference theorists, when distinguishing propositional content from cognitive significance, are reasoning in terms of *sentence types*. It is therefore not clear how New B-theorists could resort to the New Theory of Reference in their defense of *token-reflexive* tenseless truth-conditions. However, in order for the B-theorist to be in a position to criticize the A-theory, it suffices that propositional contents do not involve A-characteristics—provided that A-characteristics are not presupposed in the analysis of cognitive significance.

In this article I take truth-bearers not to be sentence tokens, but contents expressed by sentence types (cf. Footnote 7). When discussing the merits and problems of the B-theory, I wish to consider a view according to which a propositional content expressed by a temporal sentence in a given context is only about time-points and the earlier-later relation among them, not about properties such as *futurity* or *presentness*. For an explicit discussion of the B-theory from the viewpoint of sentence types, see [36]. In Sect. 19.5, two variants of this kind of B-theory are discerned, to be referred to as the B_C-theory and the B_A-theory.

In the remainder of this article I show how my framework leads to a novel perspective on the A-theory/B-theory debate. I suggest that what the A-theorist perceives as tensed truth-conditions should be understood as truth-conditions relying on intentionally individuated times, while the B-theorist's position should be reconstructed as the claim that only physically individuated times are needed when accounting for the semantics of temporal language. My view goes against the B-theory, since generally we need intentionally individuated times in addition to physically individuated times. However, I do not maintain that there are A-characteristics—peculiar properties such as *pastness* or *presentness*. The reason why we need intentionally individuated times is not ontological, but is due to the way in which epistemic agents structure a space of alternatives relevant for them. Intentionally individuated times have no ontological repercussions, and they are needed in semantics only when we must consider agents, and even then only in modal settings, i.e., when several scenarios need to be considered in order to evaluate one and the same sentence.

19.4.2 What Is a Tensed Truth-Condition?

It is not easy to pin down any exact features which, if indeed exemplified, would suffice for definitely supporting the A-theory. One of the problems with the A-theory is that not everyone agrees that the basic notions of the theory are even sufficiently well understood to allow further discussion. As the A-theory is typically sketched, reference is made to 'ontological tenses' or 'transient temporal properties' or 'tensed truth-conditions'—and the opponents of the A-theory (let alone philosophers not engaged in discussions on the philosophy of time) simply tend to find the very idea of ontological correlates of grammatical tenses deprived of sense.

If one agrees to enter the discussion about the relative merits of the A-theory and the B-theory, a critique of the A-theory can be formulated with reference to the fact that we successfully use also spatially context-dependent expressions and for instance personal pronouns without getting thereby committed to anything like 'spatial tenses' or 'ontological indexicals'; cf. [6, 27]. On the other hand, at the linguistic level it can be noted that *grammatical tenses* do not have spatial or person-bound counterparts, although in languages like English there are *indexicals* for times, places and persons alike. One should, however, remain rather skeptical as to whether the linguistic asymmetry just mentioned may have ontological repercussions of the sort the A-theorist claims to identify. For example, the future tense in English has no grammatical counterpart in Finnish. And a language like French has a rich tense system, but there just is no direct and natural way to convey in French what English-speaking philosophers call 'ontological tense' or 'tensed proposition' or 'tensed truth-condition'; one must resort to metaphors like the passage of time when attempting to clarify what one means. Observations like these shed doubts on the very meaningfulness of the A-theory of time. Is this a case where language is leading the philosopher astray?

In any event the semantic behavior of sentences that according to the A-theorist give rise to tensed contents with a tensed truth-condition is different from the behavior of sentences yielding contents that also the A-theorist considers as having a tenseless truth-condition. Before commenting on potential merits of the A-theory, we must try to reconstruct the proposal it puts forward.

19.4.3 The Notion of Truth-Condition

I begin with the notion of truth-condition. Let us clarify what is at stake by proceeding for a moment from the idealization that sentences themselves rather than their contextually determined contents are to be considered as truth-bearers. We can then simply say that whatever those contexts are in which the relevant sentence types are evaluated, *truth-conditions* of those sentence types are *sets of such contexts*. More specifically, the truth-condition of sentence S would be the set of those contexts in which S is true. It is very important to be clear at this point. Truth-conditions are utterly non-linguistic entities. They must absolutely not be confused with sentences of this or that metalanguage. Truth-conditions are non-linguistic entities captured by sentences, like persons are non-linguistic entities denoted by proper names. Speaking of certain sets of contexts as truth-conditions of course suggests a link to a language, but intrinsically those sets of contexts are what they are—totally independently of a language—like persons are what they are whether or not they are named. Actually, for systematic purposes it is beneficial to resist certain connotations that naturally suggest themselves, and if $\mathcal{C}$ is a fixed set of contexts to which we have chosen to confine our attention, simply call *any* subset of $\mathcal{C}$ a *truth-condition*. Thus, a given truth-condition may or may not be captured in terms of a given language; which truth-conditions can be captured depends on the expressive resources of the language considered.

In reality the semantics of many sentences leaves various things for the context of utterance to determine (due to their containing indexicals and grammatical tenses). For systematic reasons, a truth-value must be attributed to the claim made by uttering a sentence in a context—the content it expresses—rather than to the sentence itself. This distinction is made particularly clearly in the framework of two-dimensional semantics building on the work of Stalnaker and Kaplan (cf. [4, 18, 20, 56]). Due to context-dependence, the same sentence may express different contents in different contexts. The content expressed by 'I speak' varies with the person who utters it. In each case the content expressed is true in some circumstances of evaluation and false in others. If Mary utters this sentence, the corresponding content *that Mary speaks* is false in all those counterfactual scenarios in which Mary remains silent. The content expressed by 'I speak' can never be false in the utterance context itself, as anyone uttering this sentence is ipso facto speaking.

How should the notion of truth-condition be clarified in connection with semantically context-dependent sentences? These must be conditions for the truth of a certain content or 'proposition' rather than for the truth of a sentence. Once we have extracted a content from a given sentence in a fixed context, the truth-condition of the content will be the set of those scenarios in which the content is true.[20] The difficulty is to decide what sorts of things count as contents.

[20]The relation between these 'scenarios' or circumstances of evaluation on the one hand, and the contexts that serve to determine the content on the other hand, depend on the details of the analysis. In typical formulations of two-dimensional semantics, contents are evaluated on worlds, while contexts are composed of worlds and a number of parameters including the time of utterance.

19.4.4 What Could Contents Be?

For a moment, let us carry out our discussion at a pretheoretical level, ignoring the distinction between instants and times. If in world w_0 at time t_0 Mary utters ‘I will speak’, there are at least the following five candidates for the content expressed:

(a) *That Mary speaks in w_0 at a time later than t_0*
(b) *That Mary speaks at a time later than t_0*
(c) *That Mary will speak*
(d) *That at t_0 the event of Mary’s speaking has the property of* futurity
(e) *That the event of Mary’s speaking has the property of* futurity.

A content of any of these types is a function assigning a truth-value to any suitable circumstance of evaluation. In connection with contents of types (a), (b) and (d), circumstances of evaluation are worlds, whereas contents of types (c) and (e) are evaluated relative to world–time pairs. Due to its distinction between instants and times, my semantic framework will motivate a generalized notion content. I return to this issue in Sect. 19.4.6.

In systematic semantic theories of indexical expressions, for example in the framework of two-dimensional semantics, one tends to opt for contents of type (*b*): the only context-relative factor left unspecified is the world. Consequently, the truth-value of the content may vary from world to world. In what follows I refer to contents of this type as ‘propositions’. Option (*a*) would yield a content whose truth-value is invariant over the totality of worlds: content (*a*) is true in *every* world if in the particular world w_0 Mary indeed speaks at t_0. Option (*c*) would yield a temporally indeterminate content. The content expressed by uttering ‘I will speak’ would, thus construed, not be considered as a predication about the time of utterance. It would be a tensed proposition of the kind on which Prior based his tense logic.[21] Option (*d*) as distinct from option (*b*) is available only if one is prepared to accept *futurity* as a special kind of property, not reducible to (though implying) an event’s being later than a given time-point. That is, option (*d*) is an A-theoretical version of option (*b*). Similarly, option (*e*) is an A-theoretical variant of option (*c*).

Except for type (*a*), contents of all the above types are functions assigning *true* to some circumstances of evaluation and *false* to others. There are also differences between the content types.

- A content of type (*a*) is a constant function defined on worlds. The constant value of this function on any world w is *true* iff Mary speaks in w_0 at t_0.

[21] Prior [41, pp. 15–17] did not think that propositions must be temporally determinate. The tensed sentence ‘It was raining’ would in his sense express the same proposition whenever uttered. This amounts to thinking of propositions as functions from world–time pairs to truth-values, i.e., viewing them as A-intensions in the sense of Jackson [18]; see Footnote 23. What is distinctive to Prior’s view is not that propositions in his sense are not constant functions, but that the same proposition can be expressed in contexts differing in their time-components.

- A content of type (*b*) is a function defined on worlds, assigning to w the value *true* iff Mary speaks in world w at t_0. This kind of content is a C-intension[22] in the sense of Jackson [18].
- A content of type (*c*) is a function defined on world–time pairs, assigning to a context (w, t) the value *true* iff Mary speaks in world w at time t. This is a diagonal proposition in the sense of Stalnaker [56] or an A-intension[23] in the sense of Jackson [18].
- A content of type (*d*) is a function defined on worlds, assigning *true* to w iff at t_0 in w the event of Mary's speaking has the property of *futurity*. I refer to this function as the C_A-intension of 'Mary will speak', the subscript 'A' being reminiscent of 'A-theory', not 'A-intension'.
- A content of type (*e*) is a function defined on world–time pairs, assigning to a context (w, t) the value *true* iff at time t in w the event of Mary's speaking has the property of *futurity*. This function is termed the A_A-intension of 'Mary will speak'.

In cases (*a*), (*b*) and (*d*), the content is invariably about the time t_0 and the value of the function depends at most on the world of evaluation, while in cases (*c*) and (*e*) the content is about whichever time happens to be the time of evaluation, and the value of the function depends not only on the world of evaluation but also on the time of evaluation. Hence in particular (*b*) is true in world w_1 at t_1 iff in w_1 Mary speaks at a time later than t_0, and (*d*) is true in world w_1 at t_1 iff at time t_0 the event of Mary's speaking in w_1 has the property of *futurity*, while (*c*) is true in world w_1 at t_1 iff in w_1 Mary speaks at a time later than t_1, and (*e*) is true in world w_1 at t_1 iff at t_1 the event of Mary's speaking in w_1 has the property of *futurity*.

The notion of truth-condition depends on the notion of content as follows:

- The truth-condition of (*a*) is either the set of all worlds considered or else it is empty.
- The truth-condition of (*b*) is the set of worlds w such that Mary speaks in w after t_0.
- The truth-condition of (*c*) is the set of contexts (w, t) such that Mary speaks in w after t.
- The truth-condition of (*d*) is the set of worlds w such that the event of Mary's speaking has the property of *futurity* at t_0 in world w.
- The truth-condition of (*e*) is the set of contexts (w, t) such that the event of Mary's speaking has the property of *futurity* at t in world w.

The difference between the contents (*b*) and (*c*) is not manifested in the context (w_0, t_0)—or actually in any context with the time-component t_0. However, the

---

[22] Uttering a sentence S in a context $(w_0, a_1, \ldots, a_k)$ determines a content $C[S, a_1, \ldots, a_k]$. This content is a function from *possible worlds* to truth-values. The *C-intension* of S relative to $(w, a_1, \ldots, a_k)$ is the content $C[S, a_1, \ldots, a_k]$.

[23] The *A-intension* of a sentence S is a function $A[S]$ from *contexts* to truth-values, satisfying $A[S](w, a_1, \ldots, a_k) = \mathit{true}$ iff $C[S, a_1, \ldots, a_k](w) = \mathit{true}$.

difference becomes apparent as soon as we need to consider a variety of contexts, not all carrying the time-component t_0. The same remark applies to the A-theoretic variants of these contents, namely the contents (*d*) and (*e*).

The B-theorist construes contents in accordance with option (*b*), or at least this is the most straightforward way in which this view can be understood: the content of a sentence expressed in a context is determinate with respect to temporal properties and it specifies these properties in terms of tenseless temporal relations and the time of utterance. I will formulate in Sect. 19.4.6 a variant of this B-theoretical view within my semantic framework.

19.4.5 A-Determinations vs A-Positions

What is the A-theoretic view on contents, then? An A-theorist like Craig appears to think that contents of both types (*d*) and (*e*) are relevant in this connection.

When discussing the semantic difference between the grammatical present tense and the temporal indexical 'now', Craig speaks of *A-positions*. The A-theorist takes the semantic role of the present tense to be that of ascribing to an event the A-determination of *presentness*, while the semantic role of 'now' is first of all to make reference to a contextually given now-point and second of all to ascribe the A-determination of *presentness* to an event anchored to the now-point. Borrowing an example from Craig [6, p. 252], we may compare the pair of sentences

6. The largest trachodon is now laying her eggs here on the beach
7. The largest trachodon is laying her eggs here on the beach.

On Craig's A-theoretic understanding, if C_1 is the content expressed by (6) in w_0 at t_0 in year 2000 and C_2 the content (7) expresses in w_0 at the same moment t_0, 120 million years earlier C_1 was false but C_2 might have been true. This is because due to the presence of the indexical 'now' in (6), C_1 ascribes an A-position in the sense that it makes a predication about the moment t_0, whereas C_2 does not ascribe an A-position and so it takes for its temporal anchor the time relative to which it is evaluated, instead of being once and for all anchored to t_0. That is, from the A-theorist's viewpoint C_1 is a content of type (*d*) while C_2 is a content of type (*e*). Thus, both content types are relevant in connection with the A-theory, contents of the former type arising when temporal indexicals are used and contents of the latter type when they are not used. A-determinations are involved in both cases, according to the A-theorist. Using the terminology introduced above, the content C_1 would be a $\mathrm{C_A}$-intension and the content C_2 an $\mathrm{A_A}$-intension.

One need not be an A-theorist to recognize the semantic difference between present tense and 'now'. According to the B-theorist, the sentences 'It is raining now' and 'It is raining' express at t_0 the same content: *that it rains at* t_0. Yet the B-theorist's analysis yields to the complex sentences 'It will be the case that it is raining now' and 'It will be the case that it is raining' distinct contents: at t_0, the contents they express are *that at some time s later than* t_0 *it is the case that it rains at* t_0 and *that at some time s later than* t_0 *it is the case that it rains at s*, respectively.

19.4.6 Ways of Reconstructing Tensed Contents

There are different semantic views that are all reminiscent of the A-theorist's position. Confusing these views with each other could lead to defending the A-theory for wrong reasons. It is conceivable that some such confusion is indeed responsible for the convictions of people subscribing to the A-theory. I first describe three such views and then discuss their features.

View 1: One can adopt the view I have ascribed to the A-theorists above, according to which contexts of utterance are structured in a way that involves properties such as *presentness*, *futurity* and *pastness*. This amounts to thinking that the sentence 'Mary will speak' expresses that the event of Mary's speaking has the property of *futurity*. It expresses a content of type (*e*), an $\mathrm{A_A}$-intension. And an indexical sentence like 'Mary will speak tomorrow' expresses at t_0 a content of type (*d*), according to which at t_0 the event of Mary's speaking has the property of *futurity* and this event takes place during the day following the day to which t_0 belongs. It expresses a $\mathrm{C_A}$-intension. The implied A-determinations are supposedly irreducible to B-theoretic characteristics.

View 2: One might alternatively insist that the semantic account of temporal language needs contents of type (*c*) and that these contents are conceptually primary compared with the other content candidates. When contents are considered relative to a variety of contexts (modally) and not only relative to a single context, this view differs from the B-theorist's view. For example, let C_1 be the A-intension of 'Mary will speak' and let C_2 be its C-intension determined in w_0 at t_0. Thus, C_1 is a content of type (*c*) and C_2 a content of type (*b*). If $s \neq t_0$, it can well happen that C_1 is true (false) in a world w at s while C_2 is false (true) in w. Understood in this way, contents are not 'tensed' because of some peculiar properties of circumstances of evaluation. Insofar as these contents can be used to account for the semantics of tensed language, this is because the contexts which make up their truth-conditions are *centered* worlds with possibly varying time-components, instead of simply being worlds.[24] The ideas on diagonal propositions that Stalnaker puts forward in his 'Assertion' [56] can be seen as leading to a view like this.

View 3: Finally, utilizing the framework developed in this article, the following analysis of temporal contents can be proposed—different from the View 1 and the View 2 alike. We may take contents to be *temporally structured*. A content in this sense is a structure $\langle \mathfrak{t}, f \rangle$, where $\mathfrak{t}$ is a temporal world line and f is a function assigning a truth-value to every world in the modal margin of $\mathfrak{t}$. The content $\langle \mathfrak{t}, f \rangle$ itself can be seen as a partial function from world–instant pairs (w, s) to truth-values:

[24]For centered worlds, see Quine [44] and Lewis [22].

$$\langle \mathfrak{t}, f\rangle(w, s) = \begin{cases} \textit{true} \text{ if } s = \mathfrak{t}(w) \text{ and } f(w) = \textit{true} \\ \textit{false} \text{ if } s = \mathfrak{t}(w) \text{ and } f(w) = \textit{false} \\ \text{undefined otherwise.} \end{cases}$$

Structured contents $\langle \mathfrak{t}, f\rangle$ have two independent components. The time $\mathfrak{t}$ is an inter-worldly component that singles out for every relevant world w an instant belonging to the temporal domain of w. The function f, again, is a proposition (function from worlds to truth-values)—when attention is restricted to worlds belonging to the modal margin of $\mathfrak{t}$. Applying this function does not require taking into account any cross-world considerations. The sentence 'Mary will speak' determines in a context $(w_0, \mathfrak{t}_0)$ a content $\langle \mathfrak{t}, f\rangle$, where $\mathfrak{t}_0(w_0) <_{w_0} \mathfrak{t}(w_0)$, and $f(w) = \textit{true}$ iff $\mathfrak{t}$ is realized in w and Mary speaks in w at $\mathfrak{t}(w)$. The truth-condition of the content $\langle \mathfrak{t}, f\rangle$ is the pair $\langle \mathfrak{t}, V\rangle$, where V is the set of worlds in the modal margin of $\mathfrak{t}$ satisfying $f(w) = \textit{true}$. When in a content $\langle \mathfrak{t}, f\rangle$ the time $\mathfrak{t}$ is *intentionally individuated*, we obtain sorts of temporal contents which can be seen as providing a reconstruction of the A-theorist's tensed contents but which allow us to avoid the metaphysical conclusions of the View 1. A reconstruction of the B-theorist's position, again, is obtained when in such contents the time-component is taken to be *physically individuated*. Like the View 2, also this view makes a genuine difference compared with the B-theorist's position when contents are considered over a variety of contexts. For, relative to a single world w the eventual differences between an intentional and a physical temporal world line (which are function-like entities) are not manifested—given that the two types of world lines have the same realization in the world w. This way of construing 'tensed contents' is independent of the distinction between C-intensions and A-intensions. What is crucial here is that a content is viewed as being *structured*: a proposition equipped with a temporal world line. Our recourse to two components of contents in this analysis has nothing to do with the two-dimensionality familiar from two-dimensional semantics, while such two-dimensionality is crucial for the formulation of the View 2.

The View 1 is a reasonably straightforward formulation of what the A-theorists in fact claim. The View 2 identifies a semantic phenomenon which has none of the ontological repercussions of the View 1, but serves to account for at least some of the features of tensed language that one might be tempted to consider as giving support to the A-theory—and yet the View 2 does not support the A-theory as reconstructed in the View 1. For example, consider the sentence 'John believes that Mary is speaking'. The A-theorist claims that this sentence has a reading which attributes to John a present-tensed belief, namely the belief that the event of Mary's speaking has the property of *presentness* relative to the time of evaluation. The View 2 avoids resorting to dubious properties like *presentness*. Instead, it claims that the content of John's belief is represented by the A-intension of the sentence 'Mary is speaking', and the sentence 'John believes that Mary is speaking' itself means that for every context (w, t) compatible with John's belief in w_0 at t_0 we have that Mary speaks in w at t. Since distinct such compatible contexts may well have distinct time-components, this does not amount to the B-theoretic claim according to which in all worlds w compatible with John's belief, Mary speaks in w at t_0. The content of the belief is not a B-theoretic content of type (b) discussed above.

The View 2 does not succeed in divorcing itself from the B-theoretic view insofar as we only pay attention to the truth-value of a content in a single context (and when the content is determined relative to that same context). This is what happens with extensional (non-modal) sentences like 'Mary is speaking'. The truth-value of the C-intension of 'Mary is speaking' determined in w_0 at t_0 is *true* in w_0 iff its A-intension assigns *true* to the context (w_0, t_0).[25] However, this fact does not prove that the B-theorists are right. As noted above, things change when we need to consider contents as functions defined over a number of contexts. Still, the fact that things indeed change in the modal setting does not mean that the A-theorists are right. They may simply read too much into the correct observation that there are sentences which exhibit mutually different behaviors in modal settings while behaving similarly in extensional settings. Arguably the A-theorists' comments are based on *implicitly* thinking of sentences like 'Mary is speaking' in modal settings—for instance, as expressing contents of someone's beliefs. It can be suggested that this leads the A-theorists to mistake the semantic peculiarities of suitable readings of sentences like 'John believes that Mary speaks', explicable within the View 2, for cases where such properties as *presentness* are exemplified.

The View 3, formulated in terms of my framework, offers in a sense a middle way between the Views 1 and 2. The structures relative to which semantic evaluation takes place are more complex than those postulated by the View 2. This complexity does not stem from taking the internal structure of worlds to involve A-determinations (as in the View 1), but from the fact that our understanding of the semantics requires taking into account inter-worldly considerations (temporal world lines). As explained above, the View 3 involves problematizing the very notion of content. Here contents are propositions equipped with times, and the resulting truth-conditions involve two independent components: a set of worlds structured by a temporal world line.

Apart from its notion of content, the View 3 differs from the View 2 in not resorting to anything like A-intensions. Consider again the sentence 'John believes that Mary is speaking'. In my analysis, according to the relevant reading of this sentence, evaluated in w_0 at $\mathfrak{t}_0$, there is a time $\mathfrak{t}$ intentionally individuated by John such that $\mathfrak{t}_0(w_0) = \mathfrak{t}(w_0)$ and in all worlds v compatible with what John believes in w_0 at $\mathfrak{t}_0$, Mary speaks in w at $\mathfrak{t}(v)$. On the View 3, the A-theorist's tensed

[25] This phenomenon can be compared to what Kamp [19, p. 229] notes about the behavior of temporal indexicals such as 'now': an occurrence of 'now' can only be non-vacuous if it occurs within the scope of another temporal modifier. For the distinction that I am describing to become manifest, we need to consider several scenarios, and from the linguistic viewpoint the way of forcing us to consider several such scenarios is to place a tensed verb form or a temporal indexical in the syntactic scope of an expression for a modality.

propositions are reconstructed using intentionally individuated times, and tenseless propositions using physically individuated times. The contents differ only in the type of world line they utilize. Actually, since an intentionally individuated time can coincide with a physically individuated time, the contents need not differ at all. On the View 2, tensed propositions are reconstructed using centered worlds and tenseless propositions using worlds simpliciter, so the distinction used in the View 3 is more uniform than the one the View 2 employs.

The specific strength of the View 3 lies in the fact that it allows accounting for cases where a propositional attitude is temporally specific while physically speaking mistaken. If we take ‘John believes that Mary is speaking’ to have the form $\mathrm{N}_{\mathrm{John}} B_{\mathrm{John}}\, p$, then the corresponding content being true means that even though John’s belief is not necessarily in accordance with the physical time (i.e., even if $\mathit{N\!\!\!/} B_{\mathrm{John}}\, p$ fails), nevertheless there is a temporal world line $\mathfrak{p}$, intentionally individuated by John, to which John’s belief about Mary pertains, or to put it otherwise, which temporally structures his belief. In each world v compatible with John’s belief, Mary speaks at $\mathfrak{p}(v)$. This may be compared to the analysis that the View 2 offers, which would yield a set of contexts (v, t) compatible with John’s belief, with no control over how the time components of different such contexts may or may not relate to each other. Such a content of belief—a content of type (*c*)—could hardly be qualified as temporally specific; at least nothing in the semantics would force it to be so. It appears reasonable to suppose that as a matter of fact, in many cases agents’ attitudes are temporally specific. In particular, we need to be able to discern present-tense beliefs which are temporally specific but do not necessarily pertain to the current physical time.

Neither of the Views 2 and 3 postulate anything like A-determinations and each of them offers a means to account for at least some of the phenomena the A-theorist takes to involve A-determinations. As for the structures relative to which sentences are evaluated, the View 2 postulates nothing but centered worlds with their internal structure —a structure in which in particular A-determinations are not exemplified. The reason why the View 2 nevertheless succeeds in correlating tensed sentences with contents that no tenseless sentence could have is that semantic evaluation is effected relative to *centered* worlds. By allowing a certain complication at the language/world interface—accepting A-intensions as contents—the View 2 gives rise to contents with truth-conditions that we would not obtain should we stipulate that contents must be C-intensions. In the View 3, again, we do not merely have propositions evaluated world by world, but also temporal world lines that cut across a set of worlds. How to understand the nature of this complication depends on how world lines themselves are understood. The more their epistemic character is stressed, the less these complications are metaphysical. In any event the two components of contents as identified by the View 3 enrich our semantic framework. Drawing on distinctions that can be made in this enriched setting, various phenomena can be analyzed for which the A-theorist thinks A-determinations are needed.

19.4.7 Temporal Context-Dependence

Sentences purportedly making reference to A-determinations have at least one feature in common: what they express in a context partly depends on that very context. The context provides requisite supplementary information without which the sentence could not express a definite content.

For clarity, let us note what the context-dependence of a sentence S does *not* mean. It does not mean that the *truth-value* of the content that S expresses must admit of variation. If one adopts the Priorean view on tensed sentences—the option (c) discussed in Sect. 19.4.4—it does not even require that the *content* that S expresses must change according to the context. Consider the sentence 'Yesterday is earlier than the present day'. On the B-theoretic view, this sentence exemplifies context-dependence in the (strong) sense that the content it expresses varies with the referent of 'yesterday'. On May 2, 1906, the content expressed is *that May 1, 1906, is earlier than May 2, 1906*, while on May 7, 1906, it expresses the content *that May 6, 1906, is earlier than May 7, 1906*. Still, the content C_t expressed at any one time t is true in all worlds (irrespective of the time of evaluation). On the Priorean view, the sentence 'Yesterday is earlier than the present day' expresses at all times the same content and the truth-value of this content is *true* in all circumstances of evaluation (w, t), but it still manifests context-dependence in the sense that at distinct times the relevant content is made true for different reasons: on May 2, 1906, because May 1, 1906, precedes May 2, 1906, and on May 7, 1906, because May 6, 1906, precedes May 7, 1906. In connection with other sentences, context-dependence may also be manifested by a variation of truth-value: even accepting the Priorean view and holding that 'Mary will speak' expresses uniformly the same content in all contexts, the truth-value of this content depends on the time of evaluation.

In order for a sentence supposedly invoking A-determinations to give rise to a content, temporal parameters furnished by the non-linguistic context need be employed. If there is nothing more to tensed discourse than temporal context-dependence, postulating A-determinations appears unfounded indeed. Also the B-theorist must face temporal context-dependence. If this can be successfully done, any reason to accept the A-theory vanishes. It is incontestable, whatever one thinks of A-determinations, that a grammatically tensed sentence evaluated at time t expresses a certain *B-theoretic condition*. The content the sentence 'Mary will speak' expresses at t entails the existence of a time t' later than t such that Mary's speaking takes place at t'. The A-theorist cannot deny this; what he or she can do is to claim that the semantics of the content expressed involves *in addition* the A-determination of *futurity*. Given that indeed we may perfectly well make B-theoretic claims about contextually given elements, the difference in the information conveyed by sentences like

8. Mrs. Brown is not at home
9. Mrs. Brown *is* not at home on May 7, 1906,

with (8) uttered on May 7, 1906, already appears within the B-theory: it is one thing to use context-sensitive sentences to make B-theoretic claims about an instant fixed by the context, and another to utilize 'eternal' sentences to make B-theoretic claims.[26]

How, then, does the B-theory propose to analyze the difference between the two types of sentences? The B-theorist maintains that contents of tensed sentences do not involve anything like A-determinations, and holds that insofar as temporal features are concerned, they only have to do with the temporal earlier-later relation. Thus, for the B-theorist the content expressed by (8) when uttered on May 7, 1906, is the same as the content expressed by (9). Therefore within the B-theory the informational difference between what can be communicated by (8) on the one hand and by (9) on the other cannot be accounted for *in terms of contents and truth-conditions*. It is a matter of discussion precisely how the B-theorist should clarify this difference. John Perry [38] analyzes cognitive significance by making a distinction between belief states and propositions an agent believes in virtue of being in a given belief state. Resorting to Kaplan's distinction between character (linguistic meaning) and content [20], Perry [37] and Kaplan [21] take the cognitive significance of a 'thought' to be the character of a sentence expressing the thought, whereas the object of 'thought' in a specific context would be the corresponding content.[27] Distinct characters may yield the same content in the same context; and the same character may give rise to distinct contents in distinct contexts. Even if Kaplan and Perry dissociate propositional content and cognitive significance, still they subscribe to a semantic account of cognitive significance. It is accounted for in terms of linguistic meaning: belief states are linguistically individuated. Howard Wettstein argues against this view and proposes that one could share the basic insight of the theory of direct reference while nevertheless viewing cognitive significance as lying outside the purview of semantics. As he puts it [58, p. 202]: 'There is no reason to suppose that, in general, if we successfully uncover the institutionalized conventions governing the references of our terms, we will have captured the ways in which speakers think about their referents'. Craig [6] relies on a distinction between cognitive significance and linguistic meaning in his attempt to reconstruct the B-theorist's position.

The question remains whether there is a defensible alternative to the variants of the B-theory—an alternative which could after all explicate the distinction between sentences of types (8) and (9) simply with reference to contents and truth-conditions, without recourse to more involved notions like cognitive significance, whether the latter is construed semantically in the manner of Kaplan and Perry, or non-

[26]A sentence S is *eternal* if it expresses the same content C in all contexts and within any given world the truth-value of C remains constant over time.

[27]The character of a sentence S is a function which for every context $(w, a_1, \ldots, a_k)$ assigns a corresponding C-intension $C[S, a_1, \ldots, a_k]$ in the sense of Footnote 22. The character of 'I run' assigns to the context (w_0, John, 5-7-1906) the content *that John runs on May 7, 1906* and to the context (w_0, Mary, 12-7-1976) the content *that Mary runs on December 7, 1976*. The truth-values of these contents will, then, depend on the world in which they are evaluated.

semantically as proposed by Wettstein. The A-theorists claim to offer such an alternative. In the A-theoretic analysis, the content expressed by uttering (8) on May 7, 1906, would be distinct from the content expressed by (9), since in the former case but not in the latter the corresponding truth-condition would ascribe the property of *presentness* to the relevant event. Are there further alternatives? Given that the A-theory stands or falls with the acceptability of A-determinations, and their acceptability or even intelligibility precisely poses a problem, it becomes natural to ask whether the philosophers having postulated A-determinations have simply been victims of misjudgment when resorting to a novel type of characteristic when in reality the difference is between situated and non-situated B-theoretic claims. Or is there some other conceptual distinction, not recognized by the B-theorists, that the A-theorists have so to say misperceived and thereby erroneously attempted to clarify by speaking of A-determinations? In what follows I will illustrate how the View 3 distinguished in Sect. 19.4.6 succeeds in dealing with cases in which the A-theorists have resorted to tensed contents, 'ontological tenses' and 'transient temporal properties'. Thereby I claim to identify a suitable aspect of contents of tensed sentences which goes beyond the fact that these contents are determined contextually. It is my claim that such an aspect is found via intentionally individuated temporal world lines.

19.5 Tensed Contents Reconstructed

In what follows, I use the labels 'B_A-theory', 'B_C-theory', 'A_A-theory' and 'A_C-theory' as follows. The B_A-theory amounts to the View 2 distinguished in Sect. 19.4.6. Of the different ways of understanding contents that were discerned in Sect. 19.4.4, option (*b*) gives rise to the B_C-theory, option (*e*) to the A_A-theory and option (*d*) to the A_C-theory. As I have reconstructed the A-theory, it results in contents of type (*d*) only when applied to sentences *with* temporal indexicals. When applied to sentences *without* temporal indexicals, it leads to contents of type (*e*).

I will discuss systematically a number of examples which, according to the A-theorist, pose a problem for the B-theory and require A-determinations in their semantic analysis. After indicating that both the B_C-theory and the B_A-theory encounter problems, I show how my analysis—the View 3—succeeds in resolving the discussed problems without postulating A-determinations.

19.5.1 Simple Tensed Sentences

Let us begin by considering the sentences (10) and (11):

10. Mrs. Brown is not at home
11. Mrs. Brown *is* not at home on May 7, 1906.

Sentence (11) is tenseless. According to both variants of the A-theory (A_C, A_A) and both variants of the B-theory (B_C, B_A), the content expressed by (11) is, irrespective of its context of utterance,

12. *That Mrs. Brown* is *not at home on May 7, 1906.*

Let us first compare this content with the content these theories associate with (10) when uttered on May 7, 1906. According to the two variants of the A-theory, the content expressed by (10) in w_0 on May 7, 1906, is respectively (10-A_C) and (10-A_A):

(10-A_C) *That the event of Mrs. Brown's not being at home on May 7, 1906, has the property of* presentness

(10-A_A) *That the event of Mrs. Brown's not being at home has the property of* presentness.

These contents have the same truth-value in w_0 at May 7, 1906. Because they involve an A-determination, this truth-value may differ from the truth-value of the content (12) in w_0. This suffices to explain the cognitive difference between hearing (10) and hearing (11). However, the explanation is obtained at the expense of postulating an ontologically dubious category of properties. Let us see what happens when attention is turned to the two variants of B-theory. According to the B_C-theory, sentence (10) expresses on May 7, 1906, the content (12), whereas the B_A-theory yields a different content, namely the A-intension of (10), which might be expressed as follows:

(10-B_A) *That Mrs. Brown is not at home.*

The B_C-theory fails to account for the difference between the content of (11) and the content of (10) uttered on May 7, 1906: in both cases the content expressed is (12). The B_A-theory is more promising than the B_C-theory—it assigns to (10) uttered on May 7, 1906, a content different from (12). Still at the level of truth-values in the original context of utterance this difference is not manifested: the content (1-B_A) is true in w_0 on May 7, 1906, iff the content (12) is true in w_0.

Do these observations provide a knockdown argument against the B-theoretic understanding of temporal language in favor of the A-theoretic understanding? Things are not that simple. Any overall theory of temporal expressions must account for the cognitive difference between hearing an utterance of (10) and hearing an utterance of (11). If this can be done in a B-theoretical setting, the above considerations do not offer a sufficient reason allowing a variant of the A-theory to claim priority as a semantical analysis of temporal language. Further considerations are needed.

19.5.2 Attitude Reports

A good way of studying the capacity of a theory of temporal expressions to account for the relevant differences in cognitive significance between various contents is to study the *semantics of sentences reporting on the results of having learned such contents*.[28] Instead of letting one's psycholinguistic preconceptions to take over when one reflects on how this or that agent might react on hearing someone utter (10) or (11), let us study systematically sentences like (13) and (14):

13. John thinks that Mrs. Brown is not at home.
14. On May 7, 1906, John *thinks* that Mrs. Brown *is* at home on May 7, 1906.

Even if cognitive significance had nothing to do with the propositional contents of simple sentences such as (10) and (11), it appears rather reasonable to suppose that differences in cognitive significance of these simpler sentences should be somehow manifested in the propositional contents of the complex sentences (13) and (14).

Sentence (14) attributes to John a *de dicto* belief, whose tenseless content is *that Mrs. Brown is at home on May 7, 1906*. As for (13), if it is uttered on May 7, 1906, it reports John as having a certain belief content. Which one? The four options that were distinguished above as candidates for the content expressed by (10) all suggest themselves. Since we are interested in seeing to which extent the B-theorist can account for the cognitive significance of grammatically tensed sentences, let us look at the options offered by the B_C- and B_A-theories. As before, the B_C-theory would not fare well here either. According to it, the content of (13) uttered on May 7, 1906, would be:

(13-B_C) *That on May 7, 1906, John thinks that Mrs. Brown is not at home on May 7, 1906*,

which is the same as the content expressed by (14). However, the B_A-theory works better. It assigns to (13) the content

(13-B_A) *That John thinks that Mrs. Brown is not at home*,

construing the subordinate clause 'John thinks that Mrs. Brown is not at home' via its A-intension, so that this content is true in w_0 on May 7, 1906, iff the following holds: for every context (w, t) compatible with what John believes in w_0 on May 7, 1906, we have that Mrs. Brown fails to be at her home in w at t. Therefore the content expressed by (14) can very well be true (false) in w_0 while the content (13-B_A) is false (true) in w_0 on May 7, 1906: it can happen that among contexts compatible with John's belief in w_0 on May 7, 1906, there are contexts (w, t) whose time-component t differs from May 7, 1906. In brief, what John learns when hearing someone utter (10) on May 7, 1906, is not necessarily the content (12). Whether

[28]Wettstein, arguing that cognitive significance is not a semantic matter, takes up belief reports as a potential problem issue for his position [58, pp. 205–209]. He admits that '[a]ttitudinal embeddings seem to be a place where cognitive and semantic questions converge' [p. 205]. However, Wettstein takes belief reports to pose in any case great difficulties to systematical treatment and recommends concentrating on simple, unembedded sentences. From my viewpoint such a maneuver is unacceptable, since all crucial distinctions stemming from the behavior of world lines remain latent in the semantics of simple sentences, and only come to the fore in modal settings.

this is the case or not depends on what sorts of contexts are compatible with John's overall state of belief. In the special case that these contexts happen to agree on their time-components—the time-component being in each case May 7, 1906—the content of John's belief is indeed (12).

19.5.3 Problems with the B$_A$-Theory

The B$_A$-theory avoids postulating A-determinations and yet accounts for the cognitive difference between utterances of (10) and (11). Therefore it is, other things being equal, to be preferred over the A-theory. However, there are at least three problems with the B$_A$-theory.

First, as I have argued, cross-context identity is conceptually a complex phenomenon. The way in which the B$_A$-theory deals with contexts compatible with an agent's propositional attitudes ignores this issue and represents the sameness of a time across contexts simply by having literally the same time-component in a number of contexts. Given the analysis described in the present article, it is simply meaningless to speak of contexts (w_1, t_1) and (w_2, t_2) with $w_1 \neq w_2$ but $t_1 = t_2$. The notion of 'same time' is not an unproblematic notion in modal settings.

Second, a belief which is mistaken for its temporal aspects by physical standards, is normally nevertheless temporally specific: when John has a present-tense belief according to which Mrs. Brown is absent from her home, he may be far from knowing what time it is, and yet his belief normally pertains to what he *takes to be* the present time. On the other hand, in some cases—perhaps when John has been awake for too long or has consumed too much alcohol—it may happen that his beliefs are totally unstructured, temporally confused, and they pertain to no such specific experienced moment of time. In this latter case the set $\{\mathfrak{p} \in \mathfrak{T}^{\text{John}}_{\mathfrak{t}_0(w_0)} : \mathfrak{t}_0(w_0) = \mathfrak{p}(w_0)\}$ is empty. The B$_A$-theory has no way of making a systematic difference between such cases: on the B$_A$-theorists' premises there is nothing that can be written down in their logical language that would express that John's attitude is temporally of the one kind rather than of the other.[29] John's belief may, as it happens, allow whatever kinds of doxastic alternatives (specific or not), but in our language we lack means of expressing the difference. This problem does not reside in the structures we talk about, but in the semantics of the language used for talking about them.

Third, perhaps the most severe problem for properly distinguishing (13) and (14) is that there is nothing in the analysis provided by the B$_A$-theory that gives sense to the idea that a belief pertains to the experienced present moment. It only succeeds in distinguishing a belief that pertains to the physical present moment from a belief that

[29] In **TL** the difference can be expressed by the contrast between the formulas $\mathrm{N}_{\mathrm{John}} B_{\mathrm{John}} p$ and $(B_{\mathrm{John}} \mathrm{E}_{\mathrm{John}} p \wedge \neg \mathrm{N}_{\mathrm{John}} \neg \bot)$, where $\mathrm{E}_{\mathrm{John}} p$ is an abbreviation of $(\mathrm{P}_{\mathrm{John}} p \vee \mathrm{N}_{\mathrm{John}} p \vee \mathrm{F}_{\mathrm{John}} p)$ and $\bot$ abbreviates $(q \wedge \neg q)$ for some fixed atom q.

does not. However, not pertaining to the physical present moment is not the same as positively pertaining to the experienced present moment. It could well happen that the set of contexts compatible with what John believes in w_0 at t_0 is the same as the set of contexts compatible with John's beliefs in w_0 at a later moment t_1, and so this would leave it totally unaccounted for in what sense this one and the same set could on both occasions serve to represent what John *experiences as being presently the case*.

Given these problems with the $\mathrm{B_A}$-theory, it cannot be considered as satisfactorily dealing with the cognitive difference between what one learns when hearing (10) and what one learns upon hearing (11). Let us see, then, how my framework succeeds in analyzing the relevant difference.

19.5.4 Modes at Work

I will use **TL**-formulas to identify different readings of various natural language sentences. When a sentence S of interest contains grammatical tenses or indexicals, in its logical representation ψ_S there will occur temporal quantifiers or nominals of type 2. Relative to a fixed evaluation time, ψ_S determines, then, a specific content. According to my analysis, (10) has the readings (15) and (16):

15. И Mrs. Brown is not at home
16. N_a Mrs. Brown is not at home.

In the terminology adopted in Sect. 19.3.4, (15) is a B-theoretic and (16) an A-theoretic reading. The choice between these readings depends on whether (10) is understood as speaking of the physical present time or the present time as experienced by an agent **a**. Suppose the relevant agent **a** is John. Let $\mathfrak{t}_0$ be the physical time May 7, 1906, and let $\mathfrak{t}_0^{\mathrm{John}}$ be the intentionally individuated time that John experiences as present in w_0 on May 7, 1906. Let the modal margins of these times be respectively V and V', with $w_0 \in V \cap V'$. The contents expressed by (15) and (16) on May 7, 1906, are the structures $\langle \mathfrak{t}_0, p \rangle$ and $\langle \mathfrak{t}_0^{\mathrm{John}}, p' \rangle$, where $p(v) = \mathit{true}$ iff Mrs. Brown is not at home in v at $\mathfrak{t}_0(v)$ and $p'(v') = \mathit{true}$ iff Mrs. Brown is not at home in v' at $\mathfrak{t}_0^{\mathrm{John}}(v')$, for all $v \in V$ and $v' \in V'$. As for the sentence (11), it has the form

17. $@_{\underline{5\text{-}7\text{-}1906}}$ (Mrs. Brown is not at home).

The formula (17) expresses at any time the content that (15) expresses on the specific date of May 7, 1906, namely the content $\langle \mathfrak{t}_0, p \rangle$. Since I have assumed that temporal skeletons are actualist, whatever other differences there may be between the physical time $\mathfrak{t}_0$ and the experienced time $\mathfrak{t}_0^{\mathrm{John}}$, they coincide on w_0. Even the necessity of this modest local connection between $\mathfrak{t}_0$ and $\mathfrak{t}_0^{\mathrm{John}}$ might be called into question, but I accept to render my task of establishing a difference between the cognitive roles of (10) and (11) more difficult by supposing that this connection indeed prevails.

Given that $\mathrm{t}_0(w_0) = \mathrm{t}_0^{\mathrm{John}}(w_0)$, it follows that the content $\langle \mathrm{t}_0, p \rangle$ is true in w_0 iff Mrs. Brown is not at home at $\mathrm{t}_0(w_0)$ iff Mrs. Brown is not at home at $\mathrm{t}_0^{\mathrm{John}}(w_0)$ iff the content $\langle \mathrm{t}_0^{\mathrm{John}}, p' \rangle$ is true in w_0. This is analogous to what was noted above in relation to both variants of the B-theory. However, the fact that the truth-values of the two contents coincide at the instant $\mathrm{t}_0(w_0) = \mathrm{t}_0^{\mathrm{John}}(w_0)$ is not by itself problematic. We must account for the cognitive difference between an utterance of (11) and an utterance of (10) on May 7, 1906. Therefore we must compare $\langle \mathrm{t}_0, p \rangle$ and $\langle \mathrm{t}_0^{\mathrm{John}}, p' \rangle$ as contents of John's belief. Linguistically the difference between the two contents is reflected in semantics of the sentences (13) or 'John thinks that Mrs. Brown is not at home' and (14) or 'On May 7, 1906, John *thinks* that Mrs. Brown *is* not at home on May 7, 1906'.

Corresponding to the two readings (15) and (16) of the simple sentence (10), the complex sentence (13) has a B-theoretic reading and an A-theoretic reading:

18. *И* B_{John} Mrs. Brown is not at home.
19. $\mathrm{N}_{\mathrm{John}}$ B_{John} Mrs. Brown is not at home.

According to both formulas, the set of worlds compatible with John's beliefs is structured by a temporal world line. In the former case the world line is physically individuated and in the latter case intentionally individuated. Both formulas ascribe to John a temporal *de re* belief. In fact, according to (18) the content of John's belief is $\langle \mathrm{t}_0, p \rangle$, whereas according to (19) the content of his belief is $\langle \mathrm{t}_0^{\mathrm{John}}, p' \rangle$. Since the global cross-world behavior of $\mathrm{t}_0^{\mathrm{John}}$ may to a large extent deviate from the cross-world behavior of t_0, generally there will be worlds u, $u' \in V \cap V'$ satisfying the following condition: $\mathrm{t}_0(u) \neq \mathrm{t}_0^{\mathrm{John}}(u)$ and $\mathrm{t}_0(u') \neq \mathrm{t}_0^{\mathrm{John}}(u')$ and $p(u) = p'(u') = \mathit{true} \neq \mathit{false} = p'(u) = p(u')$. When this condition holds, the contents $\langle \mathrm{t}_0, p \rangle$ and $\langle \mathrm{t}_0^{\mathrm{John}}, p' \rangle$ are logically independent of each other: there is at least one world (namely u) in which the former is true and the latter is false, and conversely there is at least one world (namely u') in which the latter is true and the former is false. John's belief can have one content without thereby having the other.

We may note in passing that if the nominal $\underline{\text{5-7-1906}}$ of type 2 stands for t_0, the following formula (20) expresses on an arbitrary date the same content that (18) expresses on May 7, 1906:

20. $@_{\underline{\text{5-7-1906}}}$ B_{John} (Mrs. Brown is not at home).

As for the sentence (14), the content it expresses can be explicated by the formula (21):

21. $@_{\underline{\text{5-7-1906}}}$ B_{John} (Mrs. Brown is not at home $\wedge$ 5-7-1906),

where 5-7-1906 is a nominal of type 1. The content $\langle \mathrm{t}_0, p'' \rangle$ expressed by (21) on an arbitrary date is otherwise the same as the content $\langle \mathrm{t}_0, p \rangle$ expressed by (18) on May 7, 1906, except that the former content specifies, for every world v in the modal margin of t_0, that the instant $\mathrm{t}_0(v)$ is denoted in v by the nominal 5-7-1906. That is, the content $\langle \mathrm{t}_0, p'' \rangle$ entails the content $\langle \mathrm{t}_0, p \rangle$ but is not entailed by it. The presence of the nominal 5-7-1906 in (21) is needed in order for the formula to ascribe a *de dicto* belief to John, i.e., it is essential for expressing that the content of John's belief is *that Mrs. Brown is not at home on May 7, 1906*. Since the contents $\langle t_0, p \rangle$ and $\langle t_0^{John}, p' \rangle$ are mutually independent, a fortiori the content $\langle t_0, p'' \rangle$ is independent of the content $\langle t_0^{John}, p' \rangle$. This accounts for the cognitive difference between what one learns when hearing (10) and what one learns when hearing (11) on May 7, 1906. The difference is manifested at the level of *contents*, though it does not become apparent if attention is confined to the actual world, as is the case when evaluating the simple sentences (11) and (10) on May 7, 1906. The difference becomes evident only when these contents are considered as contents of someone's propositional attitudes. The difference is seen when comparing the content (21) expressed by (14) on the one hand and the content expressed by the reading (19) of (13) on May 7, 1906, on the other hand.

I conclude that the semantic theory developed in this paper offers a novel perspective on the A-theory/B-theory debate. My analysis does not postulate anything like A-determinations, and it succeeds to respond to the problem of cognitive significance while avoiding the three problems of the B_A-theory discerned in Sect. 19.5.3. The cross-world notion of time is taken seriously. The *temporal specificity* of temporal beliefs—the fact that they normally pertain to a fixed time—can be expressed also in relation to times experienced by an agent. Finally, my analysis gives sense to the idea of present-tense belief as opposed to a belief characterizable negatively as not pertaining to the current physical moment. These results are obtained because my semantic framework distinguishes physically individuated times from times experienced as present by an agent.

We have obtained an explication of the cognitive difference between (10) and (11) at the level of truth-conditions of the belief reports (13) and (14). Not only is the account semantic, but more specifically it does not resort to anything like Kaplanian characters but stays at the level of propositional content. All that is required is to have recognized the two components of propositional contents: a proposition and a temporal world line. Temporal world lines as inter-worldly components of a propositional content provide a sort of conceptual ingredient that Perry in his classic paper 'The Problem of the Essential Indexical' [38] remarked to be missing in the account of propositional content to which he referred to as the 'doctrine of propositions'. My analysis is consistent with Wettstein's remark according to which identifying the institutionalized conventions that govern the semantic values of linguistic expressions does not amount to capturing the ways in which language users think about those semantic values. First, by recognizing intentionally individuated temporal world lines as constituents of certain propositional contents, the ways in which agents think about times become a matter of semantics, notably in connection with attitude reports. This is why differences in cognitive significance can be manifested at the level of truth-conditions. Second, this does not mean, of course, that in order to understand a sentence like (13), we must be able pinpoint a temporal world line intentionally individuated by John. The institutionalized conventions regarding the use of grammatical tenses tell us that they can be used to talk about times intentionally individuated by language users. These conventions do not provide us with specific world lines, any more than the convention governing the use of the quantifier 'someone' furnishes us with a specific set of individuals.

19.5.5 Further Examples

Let us take a look at further cases which according to the A-theorist lend support to the A-theoretic viewpoint. First consider the sentence

22. In 1971 Mary believed that Nixon was president.[30]

According to the B_C-theory, its content is *that in 1971 Mary* believes *that Nixon* is *president in 1971*. This analysis is wrong, since Mary may have been totally ignorant of the then-current year when holding her belief. That is, there is no reason to think that Mary would have entertained a temporal *de dicto* belief with the content *that Nixon* is *president in 1971*. Yet this is how the B_C-theory would have it. Whatever benefits the B_A-theory may have over the B_C-theory, it suffers from the general problems diagnosed in Sect. 19.5.3.

My framework discerns two readings of the sentence (22):

23. ꟼ(1971 $\wedge$ B_{Mary}(Nixon is president))
24. ꟼ(1971 $\wedge$ N_{Mary} B_{Mary}(Nixon is president)).

What (23) expresses, say, in 2011 is that in 1971 Mary had a belief pertaining to the physical time 1971, while (24) expresses in 2011 that Mary held in 1971 a belief pertaining to the time Mary then experienced as the present time. The latter belief is not a belief about a physical time, unless Mary's experience of the present time in 1971 was perfectly accurate over the relevant modal margin. Thus, (24) serves to attribute to Mary a belief that was a sort of present-tense belief when held in 1971. This is a possible reconstruction of what the A-theorist may have in mind, though our formulation does not require postulating A-determinations.

It should be noted that what (24) expresses in 2011 is *not* the same as what (25) expresses then:

25. P_{Mary}(1971 $\wedge$ B_{Mary}(Nixon is president)).

The latter content pertains, not to the time Mary experienced as present in what (physically speaking) is 1971, but to a time she experiences as past in 2011 and whose realization in w_0 is denoted by the temporal nominal '1971' of type 1 in w_0. This is a time experienced as past which coincides locally, in w_0, with the physical time 1971. If Mary herself uttered in 2011 the English sentence

26. In 1971, I believed that Nixon was president,

the content she would thereby express would allow inferring the content expressed by (25)—under a very natural reading of the sentence (26). For, the grounds Mary can have for uttering (26) have primarily to do with her recollection in 2011 instead of what she in fact perceived as present back in 1971. We may further observe that if the content expressed by (24) holds, the content expressed by (25) may hold as well. This will be the case if Mary has a very good memory and she is able to represent

[30]Sentence (27) and its constituent (22) are discussed in [6, 47, 50].

to herself in 2011 in a precise fashion the intentionally individuated time that she experienced as present in 1971.

Let us, then, consider a more complicated example:

27. In 1971 Mary believed that Nixon was president, and today she still believes that.

It is not entirely obvious how to represent the content expressed by this sentence. If it was correct to attribute to Mary a temporal *de dicto* belief, as the B_C-theory wants to have it, the content would simply be: *that in 1971 Mary* believes *that Nixon* is *president in 1971, and in 2011 Mary still* believes *that Nixon* is *president in 1971*. However, since Mary's belief need not have been explicitly about the year 1971, this is not correct. If, again, we take the sentence as attributing to Mary a temporal *de re* belief about the physical time 1971, then the logical form of (27) could be expressed by modifying the formula (23):

28. $\P[1971 \wedge B_{\text{Mary}}(\text{Nixon is president})] \wedge B_{\text{Mary}}\ @_{\underline{1971}}(\text{Nixon is president})$.

This is a possible reading of (27), but not the most natural one, since Mary may well have held her belief in 1971 while being mistaken about the exact timing of her belief in physical terms. Simply, she need not perceive in 2011 the temporal object of her past belief as the year that in fact is the year 1971; she need not even pose herself the question about the temporal location of the time of her past belief on a public time scale. In order to suitably modify the formula (24) so that the resulting formula expresses the logical form of the relevant reading of (27), we need to be able to utilize the time introduced as the semantic value of N_{Mary} further on within the modified formula, so that we can state Mary's current belief to pertain to the very same past intentionally individuated time that witnesses the operator N_{Mary}.

Since I have chosen to use a modal-logical syntax without explicit syntactic variables, the semantic value of a tense operator is not automatically available in an arbitrary position in its syntactic scope. In order to free ourselves of this limitation, different options are available. I prefer to utilize a device familiar from hybrid logic known as the $\downarrow$-binder. I want to be able to 'store' the semantic value of an operator so that this value can be taken into use later in the evaluation. To this end, we introduce a set TVAR consisting of *time variables* $\mathrm{t}, \mathrm{t}', \ldots$, and use the binder $\downarrow$ so that the semantic effect of $\downarrow\mathrm{t}$ is to store the current time as the value of the variable t: the formula $\downarrow\mathrm{t}\phi$ holds in w at $\mathfrak{s}$ under an assignment γ over time variables iff the formula ϕ holds in w at $\mathfrak{s}$ under the assignment $\gamma[\mathrm{t}/\mathfrak{s}]$. (Here $\gamma[\mathrm{t}/\mathfrak{s}]$ is the assignment of type TVAR $\rightarrow$ **time** which is otherwise like γ except that it maps the variable t to the time $\mathfrak{s}$.) The utility of the $\downarrow$-binder lies in the fact that even if the evaluation of ϕ may lead to shifting the time of evaluation, the time $\mathfrak{s}$ remains available to be used later in the evaluation because it is stored as the value of the variable t—provided that this variable has not been reinterpreted in the course of the evaluation.[31] For example, the formula

[31]For representing contents of natural-language sentences, a sparing use of the $\downarrow$-binder suffices. As a matter of empirical fact about English, we never need to access arbitrary times introduced in the course of evaluating a discourse. Cf. the discussion on interpretive dependence in [17].

29. It is raining $\rightarrow$ $\downarrow$t ꟻ $@_{\mathrm{t}}$ it is raining

expresses in a world w_0 at a time t_0 that if it is raining in w_0 at $\mathrm{t}_0(w_0)$, it will be the case that it is raining in w_0 at $\mathrm{t}_0(w_0)$. In other words, (29) expresses the same as the formula

30. It is raining $\rightarrow$ ꟻ $@_{\underline{\mathrm{now}}}$ it is raining.

The following is an example in which the $\downarrow$-binder is not eliminable in favor of satisfaction operators and temporal indexicals:

31. ꟼ(It is raining $\wedge$ $\downarrow$t $@_{\underline{\mathrm{now}}}$ ꟻ $@_{\mathrm{t}}$ it is raining).

This formula explicates the logical form of the sentence 'It was raining, and it will be the case that it was raining then'.

With our formalism extended by the $\downarrow$-binder, the logical form of the most natural reading of (27) can be represented as follows:

32. ꟼ$\big[$1971 $\wedge$ $\mathrm{N}_{\mathrm{Mary}}$ $\downarrow$t $\big(B_{\mathrm{Mary}}$(Nixon is president) $\wedge$ $@_{\underline{\mathrm{now}}}$ B_{Mary} $@_{\mathrm{t}}$(Nixon is president)$\big)\big]$.

The belief ascribed by the formula (32) to Mary in 2011 is a temporal belief structured according to the very same time that structured her belief in 1971, namely the intentionally individuated time Mary experienced as present in 1971. This is the most faithful rendering of the use of the anaphoric 'still believes that' construction.[32]

For a further illustration of how my framework deals with cases the A-theorists have taken to favor their viewpoint—cases particularly problematic for the $\mathrm{B_C}$-theorists—let us consider Frege cases for times: attitude reports involving temporal expressions co-referential in the actual world.[33] Let us consider the sentence

33. It is now 4:30.

The content that (33) expresses at 4:30 is, for the $\mathrm{B_C}$-theorist, the tautologous content *that it is 4:30 at 4:30.* Yet uttering (33) at 4:30 can of course be very informative indeed. Following the methodology I adopted in Sect. 19.5.2 for studying the cognitive significance of contents of simple sentences, let us consider the content expressed by the following attitude report at 4:30:

34. John knows that it is now 4:30.

According to Stalnaker [56], what a speaker wishes to communicate by uttering a sentence such as (33) is a diagonal proposition (i.e., an A-intension). His view leads to construing (34) in accordance to the $\mathrm{B_A}$-theory. While this way of understanding

[32]Without affecting the truth-condition, we could change the syntactic position of $\downarrow$t, and place it between the second occurrence of the conjunction symbol and the satisfaction operator $@_{\underline{\mathrm{now}}}$.

[33]For temporal Frege cases of this kind, see [6, p. 257], [7, pp. 77–80], [27, Ch. 5], [31, 54, 56]. For Frege's general discussion on how the choice between co-referential expressions may change the cognitive value of a sentence, see [9].

the contribution of (33) on an agent's cognitive state is of course a considerable improvement compared to what the B_C-theory would predict (namely that the agent learns a tautology), this solution is not satisfactory—due to the general problems with the B_A-theory that were diagnosed in Sect. 19.5.3.

In my framework, the logical form of the sentence (34) can be understood as being (35) or (36), depending on whether the indexical 'now' is construed as referring to the physical time of utterance (indeed 4:30) or rather the present as experienced by John:

35. $@_{\underline{\text{now}}}\, K_{\text{John}}\, 4{:}30$
36. $@_{\underline{\text{now(John)}}}\, K_{\text{John}}\, 4{:}30$.

Neither the content expressed by (35) at 4:30 nor the content expressed by (36) at 4:30 is tautologous. Let $\mathfrak{t}_0$ be the physical time 4:30. Let us consider (35) first. The content this formula expresses at 4:30 is false if John does not know which time 4:30 is, in the sense that there are worlds v compatible with his knowledge such that the nominal '4:30' of type 1 does not refer to the realization $\mathfrak{t}_0(v)$ of the physical time $\mathfrak{t}_0$ in world v: $\mathfrak{t}_0(v) \neq \mathbf{Val}(4{:}30, v)$. Actually, (35) expresses that John knows the intended use of the temporal expression '4:30': (35) is true at 4:30 precisely when the physically individuated time to which 'now' refers indeed is realized in all worlds compatible with John's knowledge as the denotation of the nominal '4:30'. That is, uttering at 4:30 the sentence (34) in the sense (35) simply expresses that John knows how the expression '4:30' is normally used. This goes in the direction of the B_C-theory: understood in this way, what the sentence (34) expresses at 4:30 does not in any way link the content of John's knowledge to what he experiences as present. Yet according to our analysis the knowledge ascribed to John is not vacuous, since John might fail to use the expression in the intended way, i.e., so as to satisfy $\mathfrak{t}_0(v) = \mathbf{Val}(4{:}30, v)$ for all worlds v in the relevant modal margin. The knowledge thereby ascribed to John by (35) is essentially of linguistic nature, pertaining to the way in which the nominal '4:30' is used.

Now, sentence (34) is normally not used to convey (35) but (36). Suppose John is familiar with the intended linguistic convention regulating the use of the expression '4:30', i.e., suppose (35) holds at 4:30. Under this assumption, (36) expresses at 4:30 a relationship between the current physical time ($\mathfrak{t}_0$) and the present time as experienced by John ($\mathfrak{p}$). The former is a physically individuated temporal world line, the latter being intentionally individuated. Generally, there is no reason to expect that $\mathfrak{t}_0$ and $\mathfrak{p}$ coincide. The non-trivial information that John may acquire when a sufficiently reliable source utters (33) at 4:30 is that the time he experiences as present indeed coincides with the present physical time: that over all worlds v compatible with John's knowledge in w_0 at $\mathfrak{p}$, we have $\mathfrak{t}_0(v) = \mathbf{Val}(4{:}30, v) = \mathfrak{p}(v)$.

The following formula expresses, at any time at which it is evaluated, that (in a non-trivial sense) John knows what time it is:

37. $@_{\underline{\text{now(John)}}}\, K_{\text{John}}\, \underline{\text{now}}$.

Namely, if $\mathfrak{p}$ is the time John experiences as present, the content expressed by (37) in w_0 at a contextually given physically individuated time $\mathfrak{t}_0$ is that $\mathfrak{t}_0(v) = \mathfrak{p}(v)$, for all worlds v compatible with John's knowledge in w_0 at $\mathfrak{t}_0$. It appears reasonable to take knowing what time it is to mean that the time one experiences as present captures the actual physical time.

For a final illustration of analyses my framework offers, consider (38) and (39):

38. Mary's birthday is tomorrow
39. Mary's birthday is today.

When discussing this pair of sentences, the A-theorist Craig notes that in a sense it is hard to deny that if someone believes (38) on one day and (39) on the following day, he or she believes the very same content on the two occasions [6, p. 255]. He adds that it would seem reasonable to think that the very possibility of communicating information across time requires that no information be lost in the transition from the former belief to the latter. Such considerations lead Craig to accept for the sake of discussion that propositional contents are, after all, tenseless. He then seeks to reformulate his argument for the indispensability of tensed facts, proposing that either we must take the 'mode of presentation' of tenseless contents to be tensed, or postulate non-propositional tensed facts; see [6, pp. 255–264], [7, pp. 108–129]. Now, my framework can *both* accommodate a sense in which the two sentences mentioned express the same information content *and* a sense in which they express different information contents. The analysis relies once again on the contrast between physically and intentionally individuated temporal world lines.

I take the *information content* of a sentence such as (38) or (39) to be the content of the reading that involves only physically individuated times. The relevant readings of (38) and (39) have the respective forms (40) and (41):

40. $@_{\underline{\text{tomorrow}}}\, q$
41. $@_{\underline{\text{today}}}\, q$,

where q is a propositional atom true of exactly those instants that belong to one of Mary's birthdays. For the sake of facilitating discussion, let us ignore the fact that the nominals '<u>tomorrow</u>' and '<u>today</u>' (both of type 2) should actually be taken to denote intervals. Suppose that (38) is uttered at noon on May 6, 1906 (context k_1 with the physical now-point $\mathfrak{t}_1$) and that (39) is uttered at noon on May 7, 1906 (context k_2 with the physical now-point $\mathfrak{t}_2$). In k_1, the nominal '<u>tomorrow</u>' denotes the time $\mathfrak{t}_2$, and in k_2 the nominal '<u>today</u>' denotes the same time $\mathfrak{t}_2$. In both contexts the physical time $\mathfrak{t}_2$ is available as a temporal parameter, in the latter case as the first time parameter (the time of utterance), in the former as the denotation of the indexical term '<u>tomorrow</u>'. Consequently, the content expressed by (40) in w_0 at $\mathfrak{t}_1$ is exactly the same as the content expressed by (41) in w_0 at $\mathfrak{t}_2$, namely the pair $\langle \mathfrak{t}_2, p \rangle$, where the proposition p satisfies $p(v) = \mathit{true}$ iff $\mathfrak{t}_2(v) \in \mathbf{Val}(q, v)$, for all worlds v in the modal margin of $\mathfrak{t}_2$. Such a content is worthy of being termed 'information content'. Considering the belief reports

42. $B_{\text{Mary}}\ @_{\underline{\text{tomorrow}}}\, q$
43. $B_{\text{Mary}}\ @_{\underline{\text{today}}}\, q$,

the former uttered in w_0 at $\mathfrak{t}_1$ and the latter in w_0 at $\mathfrak{t}_2$, we may observe that the belief ascribed to Mary is the same in the two cases.

On the other hand, the sentences (38) and (39) can also be read differently:

44. $@_{\underline{\text{tomorrow(a)}}}\, q$
45. $@_{\underline{\text{today(a)}}}\, q$,

where 'a' indicates an agent by whom the relevant times are intentionally individuated. Here reasonable candidates are Mary herself, as well as the utterer and the addressee of the sentences (38) and (39). Let $\mathfrak{p}_1$ and $\mathfrak{p}_2$ be the intentionally individuated times respectively denoted by '$\underline{\text{tomorrow(a)}}$' in k_1 and '$\underline{\text{today(a)}}$' in k_2. The overall modal behavior of what an agent anticipates as the following day on May 6, 1906, may well differ from the modal behavior of what the same agent experiences as the present day on May 7, 1906. Indeed, it can very well happen that $\mathfrak{p}_1 \neq \mathfrak{p}_2$. Furthermore, neither of these two intentionally individuated times needs to coincide with what physically speaking is the present day on May 7, 1906: both $\mathfrak{p}_1$ and $\mathfrak{p}_2$ may well be distinct from the physically individuated time $\mathfrak{t}_2$. Generally speaking, the content expressed by (44) in w_0 at $\mathfrak{t}_1$ is not the same as the content expressed by (45) in w_0 at $\mathfrak{t}_2$. The fact that these contents involve intentionally individuated times can be considered as a reason for not counting them as *information contents*. As suggested above, we may take precisely those propositional contents to be information contents that only make use of physically individuated times. Consider, then, the attributions of belief to Mary in the case that the agent **a** is Mary herself, namely (46) and (47):

46. $B_{\text{Mary}}\, @_{\underline{\text{tomorrow(Mary)}}}\, q$
47. $B_{\text{Mary}}\, @_{\underline{\text{today(Mary)}}}\, q$,

the former uttered in w_0 at $\mathfrak{t}_1$ and the latter in w_0 at $\mathfrak{t}_2$. The contents Mary believes on the two occasions are respectively $\langle \mathfrak{p}_1, p' \rangle$ and $\langle \mathfrak{p}_2, p'' \rangle$, where the propositions p' and p'' satisfy $p'(v') = \textit{true}$ iff $\mathfrak{p}_1(v') \in \mathbf{Val}(q, v')$ and $p''(v'') = \textit{true}$ iff $\mathfrak{p}_2(v'') \in \mathbf{Val}(q, v'')$, for all worlds v' in which $\mathfrak{p}_1$ is realized and all worlds v'' in which $\mathfrak{p}_2$ is realized. The contents $\langle \mathfrak{p}_1, p' \rangle$ and $\langle \mathfrak{p}_2, p'' \rangle$ can well be different: the former is the content of Mary's belief in k_1, temporally structured via the way Mary anticipates what then is the following day, and the latter is her content of belief in k_2, temporally structured through the way in which she experiences the then-current day. Whether a difference of a relevant magnitude emerges between the two contents depends on what Mary's expectations and perceptions are at the different times. It also depends on which worlds are to be taken into account as compatible with her belief. What is of interest to the analysis of tensed language is that generally (46) and (47) are logically independent of each other, unlike (42) and (43).

19.6 Conclusion

I formulated a framework in which the notion of cross-world identity is problematized in modal settings, i.e., in connection with conceptualizations that require considering several mutually incompatible scenarios. This starting point led me to propose a novel semantic analysis of temporal quantifiers, useful for discussing grammatical tenses and temporal indexicals in natural languages. From the viewpoint of the framework I put forward, times themselves must be seen as *world lines* when they need to be considered in relation to a variety of possible worlds. Just like the notion of 'same individual' is problematic in many-world settings, so is the notion of 'same time'.

In my framework, temporality is *not* analyzed as yet another modality, as in typical formulations of temporal logic. True, my temporal quantifiers semantically resemble modal operators by being bounded quantifiers: they range over times bearing a specified relation to the current time of evaluation. However, whereas modal operators range over worlds, my temporal quantifiers are of a higher logical type and range over function-like entities taking possible worlds as arguments.

My framework was shown to offer a fresh perspective on the A-theory/B-theory debate. I reformulated the notion of content: in my sense, contents are structured. They are pairs $\langle \mathfrak{t}, p \rangle$, where p is a proposition and $\mathfrak{t}$ is a temporal world line. I pointed out that the distinction between physically and intentionally individuated times allows making a number of distinctions that the A-theorists have taken to support their view. My analysis avoids, however, the A-theorist's metaphysical conclusions: we are not committed to any peculiar characteristics such as A-determinations, ontological correlates of grammatical tenses. What is more, my view also goes against the B-theory, in holding that there are tensed contents—if by 'tensed contents' we mean contents $\langle \mathfrak{t}, p \rangle$ structured by an intentionally rather than physically individuated time. Further, my analysis allows explicating the difference in cognitive significance between tensed sentences and the corresponding tenseless sentences at the level of contents and their truth-conditions, without recourse to Kaplanian characters or belief states as distinguished from believed propositions.

Since the crucial distinction between physically and intentionally individuated times can only become apparent in the presence of agents (there can be no intentionally individuated times without agents), the semantic difference between tensed and tenseless sentences that my view recognizes is ultimately based on epistemological considerations. This is so irrespective of how much metaphysical weight one assigns to physically individuated times.

References

1. Anscombe, G.E.M. 1965. The intentionality of sensation: A grammatical feature. In *Analytical philosophy*, ed. R.J. Butler, 2nd series, 158–180. Oxford: Blackwell.
2. Areces, C., and B. ten Cate. 2007. Hybrid logics. In *Handbook of modal logic*, ed. P. Blackburn, J. van Benthem, and F. Wolter, 821–868. Amsterdam: Elsevier.
3. Broad, C.D. 1938. *An examination of McTaggart's philosophy*, vol. 2. Cambridge: Cambridge University Press.
4. Chalmers, D. 1996. *The conscious mind*. New York: Oxford University Press.
5. Chisholm, R.M. 1967. Identity through possible worlds: Some questions. *Noûs* 1(1): 1–8.
6. Craig, W.L. 1996. The new B-theory's *tu quoque* argument. *Synthese* 107(2): 249–269.

7. Craig, W.L. 2000. *The tensed theory of time: A critical examination*. Dordrecht: Kluwer Academic.
8. Dyke, H. 2002. Tokens, dates and tenseless truth conditions. *Synthese* 131(3): 329–351.
9. Frege, G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50.
10. Frege, G. 1918. Der Gedanke. Eine Logische Untersuchung. *Beiträge zur Philosophie des deutschen Idealismus, (1918–1919)* I: 58–77.
11. Geach, P.T. 1979. *Truth, love and immortality: An introduction to McTaggart's philosophy*. London: Hutchinson.
12. Hawley, K. 2001. *How things persist*. Oxford: Clarendon.
13. Hintikka, J. 1969. *Models for modalities*. Dordrecht: Reidel.
14. Hintikka, J. 1975. *The intentions of intentionality and other new models for modalities*. Dordrecht: Reidel.
15. Hintikka, J. 2006. Intellectual autobiography and replies to the contributors. In *The philosophy of Jaakko Hintikka*, The library of living philosophers, vol. 30, ed. R.E. Auxier and L.E. Hahn. Chicago: Open Court.
16. Hintikka, J., and G. Sandu. 1995. The fallacies of the New Theory of Reference. *Synthese* 104(2): 245–283.
17. Hornstein, N. 1990. *As time goes by: Tense and universal grammar*. Cambridge, MA: MIT.
18. Jackson, F. 1998. *From metaphysics to ethics: A defence of conceptual analysis*. Oxford: Oxford University Press.
19. Kamp, H. 1971. Formal properties of 'now'. *Theoria* 37(3): 227–273.
20. Kaplan, D. 1979. On the logic of demonstratives. *Journal of Philosophical Logic* 8(1): 81–98.
21. Kaplan, D. 1989. Demonstratives. An essay in the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein, 481–564. Oxford: Oxford University Press.
22. Lewis, D.K. 1979. Attitudes *de dicto* and *de se*. *Philosophical Review* 88(4): 513–543.
23. Lewis, D.K. 1983. *Philosophical papers*, vol. 1. Oxford: Oxford University Press.
24. Lewis, D.K. 1986. *Philosophical papers*, vol. 2. Oxford: Oxford University Press.
25. Lewis, D.K. 1986. *On the plurality of worlds*. Oxford: Blackwell.
26. McTaggart, J. 1908. The unreality of time. *Mind* 17(68): 457–474.
27. Mellor, D.H. 1981. *Real time*. Cambridge: Cambridge University Press.
28. Mellor, D.H. 1998. *Real time II*. London: Routledge.
29. Mozersky, J.M. 2000. Tense and temporal semantics. *Synthese* 124(2): 257–279.
30. Mozersky, J.M. 2001. Smith on times and tokens. *Synthese* 129(3): 405–411.
31. Oaklander, L.N. 1990. The New tenseless theory of time: A reply to Smith. *Philosophical Studies* 58(3): 287–292.
32. Oaklander, L.N. 1991. A defense of the New tenseless theory of time. *The Philosophical Quarterly* 41(162): 26–38.
33. Oaklander, L.N. 1996. McTaggart's paradox and Smith's tensed theory of time. *Synthese* 107(2): 205–221.
34. Oaklander, L.N. 2008. Be careful what you wish for: A reply to Craig. *Philosophy and Phenomenological Research* 76(1): 156–163.
35. Oaklander, L.N. 2012. A-, B- and R-theories of time: A debate. In *The future of the philosophy of time*, ed. A. Bardon, 1–24. New York: Routledge.
36. Paul, L.A. 1997. Truth conditions of tensed sentence types. *Synthese* 111(1): 53–71.
37. Perry, J. 1977. Frege on demonstratives. *Philosophical Review* 86(4): 474–497.
38. Perry, J. 1979. The problem of the essential indexical. *Noûs* 13(1): 3–21.
39. Plantinga, A. 1974. *The nature of necessity*. Oxford: Oxford University Press.
40. Prior, A. 1957. *Time and modality*. Oxford: Clarendon.
41. Prior, A. 1967. *Past, present and future*. Oxford: Clarendon.
42. Quine, W.V.O. 1953. Reference and modality. In *From a logical point of view*, 139–159. Cambridge, MA: Harvard University Press.
43. Quine, W.V.O. 1960. *Word and object*. Cambridge, MA: MIT.

44. Quine, W.V.O. 1969. Propositional objects. In *Ontological relativity and other essays*, 139–160. New York: Columbia University Press.
45. Rebuschi, M. 2013. *Etudes logiques et philosophiques de la rationalité dans l'interaction*, vol. 3. Questions d'attitudes—Essai de philosophie formelle sur l'intentionnalité, habilitation thesis, University of Lille 3.
46. Reichenbach, H. 1947. *Elements of symbolic logic*. New York: Macmillan.
47. Richard, M. 1981. Temporalism and eternalism. *Philosophical Studies* 39(1): 1–13.
48. Russell, B. 1906. Review of *Symbolic logic and its applications* by Hugh MacColl. *Mind* 15(58): 255–260.
49. Russell, B. 1940. *An inquiry into meaning and truth*. London: Allen & Unwin.
50. Salmon, N. 1989. Tense and singular propositions. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein, 331–392. Oxford: Oxford University Press.
51. Sider, T. 1996. All the world's a stage. *Australasian Journal of Philosophy* 74(3): 433–453.
52. Smart, J.J.C. 1962. 'Tensed statements': A comment. *The Philosophical Quarterly* 12(48): 264–265.
53. Smart, J.J.C. 1963. *Philosophy and scientific realism*. London: Routledge & Kegan Paul.
54. Smith, Q. 1987. Problems with the New tenseless theory of time. *Philosophical Studies* 52(3): 371–392.
55. Smith, Q. 1993. *Language and time*. Oxford: Oxford University Press.
56. Stalnaker, R.C. 1978. Assertion. *Syntax and Semantics* 9: 315–332.
57. Tulenheimo, T. 2009. Remarks on individuals in modal contexts. *Revue Internationale de Philosophie* 63(250): 383–394.
58. Wettstein, H. 1986. Has semantics rested on a mistake? *Journal of Philosophy* 83(4): 185–209.

Chapter 20
What's So Bad About Second-Order Logic?

Jason Turner

Abstract Second-order logic is generally thought problematic by the philosophical populace. Philosophers of mathematics and logic may have sophisticated reasons for rejecting second-order logic, but ask the average philosopher-on-the-street what's wrong with second-order logic and they will probably mumble something about Quine, ontological commitment, and set theory in sheep's clothing. In this paper, I try to get more precise about exactly what might be behind these mumblings. I offer four potential arguments against second-order logic and consider several lines of response to each. Two arguments target the coherence of second-order quantification generally, and stem from concerns about ontological commitment. The other two target the expressive power of 'full' (as opposed to 'Henkin') second-order logic, and give content to the concern that second-order logic is in fact "set theory in sheep's clothing". My aim is to understand the dialectic, not take sides; still, second-order logic comes through looking more promising than we might have initially thought.

According to its detractors, second-order logic is 'not logic'. Philosophical orthodoxy seems to side with the detractors, even if it's not quite clear what their complaint amounts to. By and large, contemporary philosophers tend to regard second-order logic with suspicion, or worse.

Among bona fide philosophers of mathematics and logic the debate is considerably refined, and often hinges on points of especial interest to those disciplines. But second-order logic's bad name among the general philosophical populace isn't

Thanks to Aaron Cotnoir, Daniel Elstein, Robbie Williams, Stephen Yablo, and an audience at the University of Leeds Centre for Metaphysics and Mind for helpful comments and discussion.

J. Turner (✉)
Department of Philosophy, Saint Louis University, Adorjan Hall 3800, Lindell Blvd Suite 130, St. Louis, MO, 63108, USA
e-mail: turnerjt@slu.edu

© Springer International Publishing Switzerland 2015

A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_20

thanks to these high-level debates. No, most philosophers think second-order logic is bad thanks to some stuff Quine said about ontological commitment and its being set theory in sheep's clothing.

I am come neither to praise second-order logic nor to bury it. Rather, I want to get to the bottom of what these complaints might be, and whether they are any good. The goal isn't Quine exegesis—I'm not particularly concerned with whether Quine's overall philosophy of logic gave him good reason to reject second-order logic. The question instead is whether either of these Quine-inspired themes gives us conclusive reason to reject second-order quantification. I claim neither neutrality nor completeness; my biases on both fronts will be in plain view. Still, my hope is not to persuade you of second-order logic's virtues, but just to give you a good feel for the relevant issues and considerations.

20.1 Logic

Second-order logic's detractors claim that it is 'not logic'. But what does it mean to call something 'logic'? It will be helpful to consider both a philosophical and a technical answer.

20.1.1 Formal Systems

Logicians study *logical systems*. At their most basic, each such system consists of three things: (i) a language; (ii) a *syntactic calculus* for the language; and (iii) a *model theory* for the language.

A syntactic calculus is, at its heart, a procedure for granting some sort of 'good' status to arguments based solely on the syntactic shape of their premises and conclusions. For a given system *S*, when *S*'s syntactic calculus regards an argument from a set of premises Δ to a conclusion ϕ as good we write '$\Delta \vdash_S \phi$'. The $\vdash_S$ relation is *S*'s *proof-theoretic consequence* relation.

A model theory specifies a class of objects—usually, set-theoretic ones—called *models*, and defines a 'true in' relation that holds between sentences of the language and these models. If a sentence ϕ is true in a model M, we say that M is a *model of* ϕ, and if all the sentences in a set Δ are true on M, then it is a model of Δ. When every model of Δ in a given system S is also a model of ϕ in that system, we write $\Delta \vDash_S \phi$. The $\vDash_S$ relation is *S*'s *model-theoretic consequence* relation.[1]

[1] This can be generalized in various ways. For instance, in many-valued logics, models assign sentences one of several truth-values, one or more of those values is 'designated', and model-theoretic consequence is understood as preservation-of-desigation-in-all-models. Other generalizations are possible, of course. For our purposes, though, we can stick with our first-pass understanding.

If a given system is *sound*, then proof-theoretic consequence in it guarantees model-theoretic consequence in it. If it is *complete*, then model-theoretic consequence guarantees the proof-theoretic kind. If we're lucky, we will have a sound and complete system. Occasionally we're not so lucky. Some systems are unsound or, more commonly, incomplete.

20.1.2 Genuine Logic

Formal systems, and the consequence relations they give rise to, are cheap. I can cook up a formal system where ϕ is both a proof-theoretic and a model-theoretic consequence of Δ if and only if ϕ has at least one instance of the name 'Jason' in it. But such a system is manifestly uninteresting. We want to study more interesting systems than that.

We might care about a formal system for one of two reasons. First, it might be *mathematically* interesting—it might have mathematical properties worth studying in its own right. But second, it might be *philosophically* interesting, because it may give us a mathematically tractable way to study something we care about: *genuine logical consequence*.

Here's the idea. Between premises and a conclusion there can hold a particular relation (or, perhaps, one of a handful of relation*s*—see [2]) that we care about. We aim to give arguments where this relation (or one of these relations) holds between the premises and conclusions of arguments. Arguments where this happens are good, and command our attention, in ways that arguments without it aren't and don't.[2] Call this relation *logical consequence*.

Perhaps (for a given language) this relation is co-extensional with the proof- or model-theoretic consequence relation of some formal system. Perhaps (for a given language) this relation is even *identical* to some proof- or model-theoretic consequence relation. Still, in the first instance we care about the relation not under its proof- or model-theoretic guise, but under a guise that we already implicitly grasp, however darkly, in our native ability to tell what 'follows from' what.

Let's use '$\Rightarrow$' for this genuine consequence relation. Its importance stems in large part from the theoretical role it is supposed to play. That role is complex, but is often thought to include:

[2] Gillian Russell [20] argues that there is a crucial ambiguity in what we take 'arguments' to consist in which leads to a logical pluralism of a different sort than that advocated in [2]. I have slid over this ambiguity, and am going to proceed by assuming (hoping!) it won't affect any of what's to come.

Modality: If $\Delta \Rightarrow \phi$, then it is in some sense impossible for all of Δ to be true while ϕ is false.[3]

Normativity: If $\Delta \Rightarrow \phi$, then it is an error in reasoning to accept all of Δ while rejecting ϕ.

Topic Neutrality: Whether $\Delta \Rightarrow \phi$ should not depend on the truths of any particular subject-matter.

Ontological Innocence: Whether $\Delta \Rightarrow \phi$ should not depend on the existence of things of any particular kind K.

(Cf. also [2, §§ 2.3–2.5]. Note that Ontological Innocence is plausibly taken as a corollary of Topic Neutrality, since whether there are things of a particular kind K looks like a truth of a particular subject-matter.)

To say that second-order logic *is* logic is to say that there is at least one second-order system *SO* where, for any Δ and ϕ in that system, $\Delta \Rightarrow \phi$ iff $\Delta \vDash_{SO} \phi$. Conversely, to say that second-order logic is *not* logic is to deny that any second-order system coincides with genuine consequence this way.

The arguments to be considered each aim to show that every second-order system violates one of the four principles just outlined. If they're right, and if those principles are indeed constraints on genuine consequence, then no second-order consequence relation coincides with genuine consequence, and so second-order logic is not logic.

You may doubt that there *is* any relation of 'genuine consequence'. Perhaps our native grasp of 'following from' is too dark and muddled to have settled on any single, unambiguous relation between premises and conclusion. If it hasn't, then the question 'Is second-order logic?' may have no well-defined answer. But this won't rob the arguments to be considered of their force, for (if successful) they can be taken to show that second-order logic fails to have certain nice features we commonly associate with 'logic'.

20.2 Second Order Systems

As I've interpreted it, the claim that second-order logic is 'not logic' depends on certain logical systems counting as 'second-order'. I should say something about what this means.

20.2.1 Second-Order Languages

Second-orderness is, in the first instance, a property of *languages*—the property a language has when it allows variables to occur in predicate position and quantifiers

[3] These principles make free use of truth and falsity, and thus may be subject to worries of the sort Field [6] has levelled against what he calls 'The Validity Argument'. I think the situation here can be finessed with use of conditionals and infinite conjunctions, but for our purposes won't bother.

to bind those variables.[4] We get a standard second-order language by taking a first-order language, adding a new stock of predicate-variables, and allowing those variables to be bound by the quantifiers. We'll call a system second-order just in case its language is second-order.

Generally, the second-order predicate variables have a fixed adicy: there will be one-placed predicate variables 'X^1', 'Y^1', ..., two-placed predicate variables 'X^2', 'Y^2',..., and so on. If a first-order formula is well-formed, then the result of replacing any of its n-adic predicates with an n-adic predicate variable will be well-formed, as will the result of binding those variables with a universal or existential quantifier.

20.2.2 Second-Order Model Theories

Second-order systems come in two varieties, depending on what their model-theory is like. Very roughly, the two model theories correspond to different conceptions of what the second-order quantifiers are doing.

Suppose we think of second-order quantifiers as ranging over properties and relations. On an *abundant* conception of properties, every set of things corresponds to some property, and every set of n-tuples corresponds to some n-adic relation.[5] On a *sparse* conception, properties and relations are scarce: some sets of individuals (or n-tuples of individuals) may correspond to no genuine property (or relation).

Full second-order systems hardwire the abundance of properties and relations into the model theory. *Henkin* second-order systems, by contrast, allow the properties and relations to be sparse.

In standard first-order model theory, models consist of a domain—a non-empty set—and an interpretation function, which assigns individuals from that domain to names of the associated language and extensions over that domain to its predicates. (An extension for a monadic predicate is simply the set of its satisfiers; the extension for an n-adic one is the set of n-tuples that satisfy it.) A variable assignment assigns individuals from the domain to variables; truth on a model is first defined for open formula relative to a variable assignment. Truth on the model *simpliciter* is simply truth on all variable assignments.

To get a model theory for *full* second-order logic, we use the same sorts of models, but modify the truth-on relation. First, we extend variable assignments so that they also assign n-adic extensions to n-adic predicate variables. We then add a clause to our definition of truth-on-a-model so that $\ulcorner \forall X \phi(X) \urcorner$ is true exactly when

[4]Systems that allow the variables but not the binding are also possible—see e.g. [21, p. 62]—but we won't consider them here.

[5]Since properties are generally taken to be intensional entities, a better characterization would use sets of possibilia. We will stick to extensional contexts here, though, so we can safely ignore the difference.

$\ulcorner \phi(X) \urcorner$ is true on all variable assignments. This essentially gives the second-order n-adic variables the effect of quantifying over all n-adic extensions drawn from the domain.

Henkin model theory differs from full model theory by including a 'second-order domain'. We can give it one in several ways, but however we do it, the effect will be to specify a range of extensions (drawn from the first-order domain) for the second-order variables to range over. We then modify the definition of 'variable assignment' so that the second-order variables can only be assigned extensions from the second-order domain, and keep the definition of truth-on-a-model the same as it is for full second-order logic. This essentially gives the second-order n-adic variables the effect of quantifying over a privileged class of n-adic extensions drawn from the domain.

This small difference has profound knock-on effects. In particular, it gives full second-order logic incredible expressive power. One example, which we will return to later, is that it allows us to form sentences (CH) and (NCH) such that, if the continuum hypothesis is true, (CH) is true on all full second-order models, and if the continuum hypothesis is false, (NCH) is true on all full second-order models.[6] As a result, full second-order systems are essentially incomplete. There are no syntactic calculi that capture all of their model-theoretic consequences. More precisely, for every full second-order system F, there is some sentence ϕ where $\vDash_F \phi$ but $\nvdash_F \phi$.[7]

Conversely, there are many complete Henkin second-order systems. In fact, this kind of system got its name when Leon Henkin [8] proved that these systems are expressively equivalent to sorted first-order systems.

The arguments to be given in Sect. 20.3 target second-order logic *generally*, and thus aim to show that *no* second-order system is logic. The arguments in Sect. 20.4 are focused primarily on showing that full second-order logic is not logic, and leave the Henkin version untouched. It's worth keeping the complaints, and their targets, distinct.

20.3 Ontological Guilt

The first group of arguments complain that second-order logic isn't ontologically innocent, and therefore that it isn't logic. (These arguments really target second-order *quantification*, rather than second-order *logic*: the idea is that, whether logic or not, there's something untoward about quantifying in predicate position.)

[6][21, p. 105] Note that (NCH) is not simply the negation of (CH). On models with countable domains, both (CH) and (NCH) are true.

[7]This is failure of what is called 'weak completeness'—not all model-theoretic truths are theorems—and thus is stronger than the failure of what is called 'strong completeness,' which happens when not all model-theoretic consequences are proof-theoretic ones.

20.3.1 Textbook Quineanism

If asked what was wrong with second-order logic, the average philosopher-on-the-street is likely to offer an argument like the following, and attribute it to Quine:

Textbook Quineanism

(i) You are ontologically committed to something of a particular kind if and only if the bound variables in your system have to range over things of that kind to be true.
(ii) The bound variables of second-order logic have to range over something predicate-like.
(iii) Therefore, theorems of second-order logic ontologically commit you to something predicate-like.
(iv) So second-order logic is not logic.

('Predicate-like' things here include sets, properties, Fregean concepts, and anything else that could in principle be the semantic value of a predicate.)

Note first that, whatever its merits, this argument cannot be Quine's. The first premise supposes we can take an arbitrary language, check its quantifiers, and figure out what its ontological commitments are. But Quine consistently held that ontological commitments were only well-defined for *first-order* theories; to get the ontological commitments of some other sort of theory, we first had to translate it into a first-order one. Quine also famously held that, in general, there was no uniquely 'right' translation for any given theory. As a result, non-first-order theories don't give rise to unique ontological commitments.

The argument can be made more Quinean by replacing the first premise with one that says we have to translate into a first-order system and replacing the second premise with the claim that we'll thereby translate second-order quantifiers as first-order ones ranging over something predicate-like. But it's not clear this argument would be any better, because it's not clear Quine's insistence on first-order translation is well-motivated. That insistence is related to his belief that serious theorizing shouldn't use anything but extensional first-order resources, and would require us to translate not just second-order quantifiers, but modal and tense operators, generalized quantifiers, and a host of other resources before we can get down to serious theory. Most philosophers nowadays reject the Quinean demand when it comes to these other resources. If they want to insist on it for second-order quantification, they owe us an explanation for the differential treatment.[8]

But even if the textbook argument isn't Quine's, it *is* an argument, and fuels contemporary suspicion of second-order quantification. Can the friend of second-order logic say anything about it?[9]

---

[8]The argument of Sect. 20.3.2 may provide such an explanation; but then we can consider it in its own right, rather than as an adjunct to Quine's.

[9]Another question: If this argument isn't Quine's, what *is* Quine's, and is it any good? Unfortunately, the closest thing I can find to an argument in Quine is at p. 66 of his *Philosophy of Logic*, and it's a howler. I've left it out in interests of space, but had I included it I would not have said anything more (or better) against it than was said by Boolos [3, pp. 510–511].

We can distinguish two very different pictures of second-order quantification. Which forms of resistance are available will depend on which picture we are attracted to. According to one picture going back to Frege, second-order quantifiers range over a special, distinctive type of object. Let's call this the *Fregean* picture. More recently, a number of philosophers have started to endorse a *qualitative* conception, according to which second-order quantifiers shouldn't be thought of 'ranging over' anything at all. To think of the quantifiers this way is to simply misunderstand the kind of business they're in.

Friends of the Fregean picture will answer the Quinean argument by denying the ontological innocence principle driving the move from (iii) to (iv). The Fregean thinks that there is a distinctive kind of *logical* entity, and that the second-order quantifiers range over those. But according to the Fregean it's no mark against a system's logicality—or even its topic neutrality—if it commits us to things of a distinctively logical kind. So the arguments of this section won't trouble the Fregean.[10]

Friends of the qualitative conception will resist premise (ii). Considering *how* they might do this will give us a better picture of what the qualitative conception amounts to.

20.3.1.1 The Nature of Quantification

Let's start by considering an argument *for* premise (ii). It runs like this: '*To use a quantifier* is to ontologically commit. The very idea of $\ulcorner \exists X^n \phi(X^n) \urcorner$ being *not* ontologically committing simply misunderstands the nature of ontological commitment. So $\ulcorner \exists X^n \phi(X^n) \urcorner$ must ontologically commit us to something that corresponds to the bound 'X^n'. Furthermore, since the bound 'X^n' is the kind of thing that can be used predicatively, it will have to be of a distinctive 'predicate-like' kind—a set, or a Fregean concept, or what-have-you. And that's all it means to say that 'X^n', ranges over something of that kind.'

According to the qualitative conception of second-order quantification, this line of thought is mistaken. Consider the sentence

(1) Fido is brown.

What does it ontologically commit us to? Intuitively, it commits us to Fido, and to brown things, although not to any predicate-like thing of 'brownness'. Following Prior [12, p. 35], we can ask why it commits us to Fido. The answer seems to be because it uses the name 'Fido,' and it uses that name in such a way that the sentence can only be true if 'Fido' refers to something. So consideration of (1) suggests that terms in name position have the power to commit, and terms in the predicate position do not.

If that is right, then the reason

---

[10]The Fregean has other troubles, though: expressibility problems that relate to the so-called 'concept horse problem'. I cannot hope to pursue that huge literature here. I'll simply focus on the qualitative conception instead.

(2) $\exists x(x \text{ is brown})$

commits us to something brown is not because a variable is bound, but because that variable is in name position. On the other hand, since 'is brown' doesn't commit us to anything predicate-like in (1), the bound second-order variable in

(3) $\exists X^1(X^1(\text{Fido}))$

shouldn't, either. To say otherwise is to say that it is somehow the distinction between the *particular* and the *general* that engenders ontological commitment. But why should this be? Why is it instead not the distinction between the *subject* and the *qualitative way* the subject is said to be that engenders the commitment? It is the presence of expressions (names or variables) in name-like position that gives rise to ontological commitment—bound variables have nothing to do with it. When we assert (3), we're not saying that there is some predicate-like entity or other that Fido participates in, any more than when we assert (1) we say that there is a particular predicate-like entity that Fido participates in. In (1) we say how Fido is particularly; in (3) we say how he is in more generality.

20.3.1.2 Model Theory

Another argument for premise (ii) appeals to second-order model theory, as described in Sect. 20.2.2. 'Just look at that model theory. It explicitly has second-order quantifiers ranging over extensions drawn from the domain. They thus range over something predicate-like—extensions—and premise (ii) follows.'

The friend of the qualitative conception has several potential responses. First, he might sharply distinguish between a *model theory* and a *semantic theory*. The former is a set-theoretic device for mapping consequence relations. The latter gives *truth-conditions* for claims of a language. The two are, at first glance, simply different projects. Since model theory uses a notion of 'truth in a model,' we are tempted to run the two projects together. But we needn't. In the first instance, model theory provides us with a class of models and a relation between models and sentences. We could have called this relation 'zapping,' and then said that ϕ is a model-theoretic consequence of Δ whenever any model that zaps all of Δ also zaps ϕ. We may think that the thus-defined relation helps us investigate genuine logical consequence without thinking that zapping is some species of *truth*, or that zapping-on-a-model conditions gives us any particular insight into meaning or truth-conditions. If truth-on-a-model conditions have little to do with truth-conditions, then the fact that the former treat quantifiers as ranging over extensions gives us little reason to think that the latter do.

Not everyone can endorse this response. In particular, standard Tarskian accounts of logical consequence identify it with the preservation of truth on all *interpretations*—all ways of assigning meanings to the non-logical terms. They then identify interpretations with models. Among the interpretations—which are the models—is the *intended* interpretation, the one that gives genuine truth-conditions for sentences of the language. On such a picture, models *are* in the business of giving genuine truth-conditions, and—if the second-order model theory is as described in Sect. 20.2.2—those truth-conditions say that $\ulcorner \exists X^n \phi(X^n) \urcorner$ is true if and only if some set of n-tuples from the domain satisfies $\ulcorner \phi(X^n) \urcorner$.

But Tarskians friendly to second-order quantification may offer one of two other responses. First, they might say we gave the wrong model theory in Sect. 20.2.2. That model theory was given in a first-order metalanguage. But if second-order languages are kosher—as the second-order Tarskian believes they are—then their model theories should be second-order as well. Down this route lies the technical work of providing such a model theory and verifying that it works the way we want it to, a task undertaken in [17]. If successful, though, it defuses the argument, for the second-order model theory does not interpret the second-order quantifiers as first-order ones ranging over subsets of the domain.

A second response accepts the model theory of Sect. 20.2.2 but denies that it makes for unwanted ontological results. This response starts by noting that, on standard Tarskian accounts, '$F(a)$' is true if and only if the denotation of 'a' is in the extension of 'F'. Still, we don't think that simply asserting '$F(a)$' commits us to the existence of extensions. (If we did, we could argue that first-order logic isn't ontologically innocent and therefore not logic.) Although extensions show up in the official truth-conditions for '$F(a)$', they are mere theoretical apparatus of the semantics, not part of what the sentence demands *of the world*. (Cf. [16] and [26].) The sentence only demands that a be F, even if sets show up in the machinery we use for semantic theorizing. And it is what a sentence demands of the world, rather than semantic machinery, that we must take seriously when figuring out a sentence's ontological commitments.

The quantification over sets given by the model theory is, according to this response, simply more semantic machinery. If that's right, the model-theoretic treatment of the second-order quantifiers, by itself, gives us no reason to think those commit us ontologically to predicate-like entities.

20.3.2 The Intelligibility Argument

A strengthened version of the Textbook argument—one which may underlie Quine's thinking, as the seeds for it can be found in his work—stems from the thought that the ontologically innocent picture of second-order quantification just sketched is somehow unintelligible.

Here, in rough form, is the idea. Formal systems are nothing but squiggles on paper until we do something to give them meaning. But the only way we can give these expressions meaning is by explicitly defining them using expressions we already understand.[11] And there's no way to do this for second-order quantifiers

[11]Or, perhaps, by ostending their meanings; but I take this option to be unavailable for second-order quantifiers.

except by identifying them with first order quantifiers over predicate-like entities. So either the second-order quantifiers are meaningless, or else they are ontologically committal in a way that bars them from being logic.

Consider how this works for first-order quantifiers. Quine ([13, pp. 65–71] and [14, pp. 161–163]) and van Inwagen [24, pp. 18–22] tell us that sentences of the form

(4) $\exists x(\ldots x \ldots x \ldots)$

should be understood as

(5) There is something x such that $\ldots x \ldots x \ldots$.

This then tells us how to understand '∃'. Furthermore, the variables themselves should be understood as pronouns, with indices to disambiguate their anaphors. That is, (5) is understood as

(6) There is something such that ... it ... it ....

If there were extra quantifiers and pronouns, we would get confused which pronouns went with which quantifiers. We can index them with little 'x' and 'y' subscripts to avoid disambiguation; and we can write the thus-indexed pronouns simply as 'x' and 'y' rather than 'it_x' and 'it_y'. That is how to understand variables.

According to the argument, if second-order quantifiers are to be intelligible, they will need a similar story about

(7) $\exists X^n(\ldots X^n \ldots X^n \ldots)$.

The most natural way to read (7) is as

(8) There is some way X^n that $\ldots X^n \ldots X^n \ldots$,

which is simply first-order quantification over 'ways'—predicate-like things that might be identified with properties, relations, extensions, or what-have you. Thus, if second-order quantification is intelligible, it's just first-order quantification over 'ways', making it not ontologically innocent and thus not logic.[12]

The above argument essentially relies on three premises ((i)–(iii)) plus the Ontological Innocence principle (v), and runs:

The Intelligibility Argument

(i) A formal system is meaningless unless it is provided an interpretation—a specification of what its expressions mean.
(ii) We provide an interpretation by specifying the meaning of each expression using terms already understood.
(iii) The only plausible meaning for '$\exists X^n \ldots$' is 'There is a predicate-like thing X^n such that ...'.

[12] I don't know anywhere this argument is explicitly presented in this form; van Inwagen presents a similar argument against substitutional quantification in [23], and comes close to giving this one [25, p. 124]. In the latter he also ascribes something like the present argument to Quine in *Philosophy of Logic*, but I cannot quite find that argument there.

(iv) So, if second-order logic is meaningful, the second-order quantifiers say that there are predicate-like things.
(v) If the second-order quantifiers say that there are predicate-like things, then second-order logic is not logic.
(vi) Therefore, either second-order logic is meaningless, or it is not logic.

I doubt the first premise can be rejected—uninterpreted squiggles on paper are just that. There is room to resist the other two, though. We'll consider them in reverse order.

20.3.2.1 Premise (iii)

Premise (iii) may be resisted by finding a better interpretation for the second-order quantifiers than that suggested by the argument. Susan Haack [7, pp. 52–55] has suggested that we can get away from these sorts of ontological-commitment worries by interpreting the second-order quantifiers *substitutionally*.

Truth-on-a-model conditions for quantified sentences are generally given 'objectually,' saying (roughly) that $\ulcorner \exists x \phi(x) \urcorner$ is true if and only if something in the model satisfies ϕ. But substitutional truth-on-a-model conditions say instead that it is true if and only if, for some name α, $\ulcorner \phi(\alpha) \urcorner$ is true. Notice this gets away from talking about things satisfying predicates, trading only in truth. If the truth-on-the-model conditions correspond to genuine truth-conditions, we can have a sort of 'ontologically innocent' reading of the quantifier. For instance, if

(9) Zeus is a greek god

can be true without there being any Zeus, then if '∃' is read substitutionally,

(10) $\exists x$(x is a greek god)

can, too [1].

It is controversial whether (9) commits us to Zeus or not. It's much less controversial that (1) does *not* commit us to any predicate-like entity of brownness. In this case, if we interpret the second-order quantifiers substitutionally—that is, if we insist that $\ulcorner \exists X^n \phi(X^n) \urcorner$ is true if and only if, for some n-placed predicate Π^n, $\ulcorner \phi(\Pi^n) \urcorner$ is true—then we seem to have an interpretation of second-order quantification that shouldn't be understood as saying there are predicate-like things.

Well, if there really is such a thing as a 'substitutional interpretation' of a quantifier, at least. Van Inwagen [23] argues there is no such thing. My so-called 'substitutional interpretation' of the quantifier was in fact a substitutional *truth-condition*. An interpretation was supposed to be specification of *meaning*, and truth-conditions aren't obviously the same.

We could try to identify the meaning with the truth-conditions, in which case '$\exists X^n \ldots$' would mean something like 'there is an n-placed predicate which ...,' and thus would ontologically commit us to predicate-like things—namely, the predicates themselves. As van Inwagen notes, friends of substitutional quantification deny they mean *this* when they quantify substitutionally. But unless they tell us what they do mean, they haven't given us an interpretation at all.

Agustín Rayo and Stephen Yablo [18] have suggested a different, non-substitutional strategy. They start by observing that most ordinary English renderings of second-order quantification look committal because they implicitly put second-order resources in nominal position. According to them, '$\exists X^n \ldots$' shouldn't be interpreted as 'there is some way ...,' because 'some way' is a noun phrase, and foes of second-order logic will see ontological commitment wherever noun phrases are used. If an interpretation wants to be non-committal, it should de-nominalize second-order quantifiers and variables.

Consider, for instance, the second-order sentence

(11) $\exists X^2(X^2(\text{Scooby, Shaggy}))$.

If we interpret this as 'There is a way Scooby and Shaggy are related,' we treat the second-order variables as ranging over 'ways'. But Rayo and Yablo suggest we should interpret (11) as

(12) Scooby and Shaggy are somehow related.

Here, the quantifier shows up adverbially, and there's no temptation to read it as quantifying over ways—or anything else.

Rayo and Yablo use this thought to interpret second-order quantifiers in a pattern that mirrors Quine and van Inwagen's interpretation of first-order quantifiers. Second-order variables are treated as the pro-adverb 'are so related', where we use different variables for anaphoric disambiguation. A sentences of the form

(13) $\exists X^n(\ldots X^n \ldots X^n \ldots)$

is understood as

(14) Things somehow$_{X^n}$ relate such that ... are so$_{X^n}$ related ... are so$_{X^n}$ related ...

The subscripted 'X^n's just help us keep track of which 'so related's go with which 'related somehow's.[13]

Yablo and Rayo's reply seems to undercut the intelligibility argument. If we can answer van Inwagen's challenge about substitutional 'interpretations', Haack's reply will undercut it as well. But it's worth noticing that neither reply, even if it vindicates second-order quantification, clearly vindicates *full second-order logic*.

When it comes to Haack's proposal, the reason is technical. The model theory for full second-order logic, recall, has the second-order variables ranging objectually over *all* extensions that can be drawn from the domain.[14] If we trade that in for substitutional quantification, we get the effect of, at best, ranging over only what are known as the 'constructible' extensions from the domain—those extensions that

[13] 'Things somehow$_{X^n}$ relate is to be interpreted (roughly) as 'things are-or-aren't somehow$_{X^n}$ related; see [18, p. 84].

[14] I'm assuming here that the second-orderist is happy to reason instrumentally with models in this way, even if she insists that, in all seriousness, second-order variables aren't in the 'ranging over' business.

contain all and only the satisfiers of some open formula in the language. And it's well-known that ranging over only the constructible extensions is far weaker than ranging over all of them. For instance, neither of (CH) and (NCH) will be true in every model. [21, pp. 110–116]

When it comes to Yablo and Rayo's proposal, the situation is less clear. But consider

(15) $\exists X^3 \forall x \forall y \forall z (X^3(x, y, z) \leftrightarrow x = \text{Scooby} \wedge y = \text{Shaggy} \wedge z = \text{Velma})$

This will be true on all full second-order models (which assign referents to 'Scooby', 'Shaggy', and 'Velma'), because for any three things in the domain one of the potential values for 'X^3' will be the set containing just one triple with those three things. But Rayo and Yablo interpret this as

(16) Things somehow$_{X^2}$ relate such that any three things that are so$_{X^3}$ related if and only if the first is Scooby, the second Shaggy, and the third Velma.

In other words—reifying 'hows' for clarification—there is a relation had by and only by Scooby, Shaggy, and Velma, in that order. This is far from clear, though; given our grasp on adverbial quantification, it may be that however Scooby, Shaggy, and Velma are related, there will be three other things that are also so-related.[15] If so, then while Yablo and Rayo's interpretation gets us second-order quantification, it won't get us full second-order logic.[16]

20.3.2.2 Premise (ii)

The intelligibility argument offers a challenge: provide an ontologically innocent natural-language interpretation of second-order quantification. Objections to premise (iii) take up this challenge. Objections to premise (ii), in contrast, reject the challenge itself.

We should agree that, if uninterpreted, second-order resources don't mean anything. But why think that meaning can *only* be assigned by explicit definition in already-understood terms? That's a pretty restrictive demand, and would seem to rule out theoretical terms like 'superposition' from quantum physics or '∈' from set theory, as neither has any explicit definition in more familiar terms. We'd better not commit to a constraint on interpretation that makes these meaningless.

Lewis [11] suggests that terms such as these are defined by *theoretical role*. We specify our theory (quantum mechanics, say, or set theory), and in doing so we use

[15]Likewise, it may be that however Scooby, Shaggy, and Velma aren't related, there will be three other things also so unrelated; this takes care of the 'don't' part of the 'do-or-don't' clause mentioned in Footnote 13.

[16]See [19] for related worries.

some new expressions, such as 'superposition' or '∈'. The theory links these new expressions with ones we're already familiar with. For instance, set theory has the theorem

(17) $\forall x \exists y(x \in y \wedge \forall z(z \in y \leftrightarrow z = x))$,

which uses not just the new '∈' but the old '∃', '=', and so on.

Call a theory *realizable* if and only if its new expression can be interpreted in a way that—leaving the interpretations of the old expressions alone—makes the theory, or at least most of the theory, true. We can then give the new expressions meaning indirectly, by saying 'Let the new expressions be interpreted however they need to be in order for the theory to be true'. If the theory is not realizable, our new terms fail to get a meaning. If there is a uniquely best interpretation of the new terms that makes the theory true, the new terms will be thus interpreted. If several equally good interpretations each would make the theory true, the new terms will be indeterminate in interpretation among those several meanings. In any case, if the theory is realizable, we can interpret its new terms without explicit definition.

If we can do this for predicates we can presumably do it for expressions of other syntactic categories. In particular, we ought to be able to do it for quantifiers. We can write down the 'theory' of second-order logic, say 'Let the second-order resources be interpreted however they must be in order to make the theory true,' and let the interpretative chips fall where they may.

What would count as the 'theory' in this case? If we just want to quantify second-order, and are happy for Henkin consequence to govern our quantifying, the 'theory' can be the axioms of some Henkin system. But since full second-order logic is incomplete, no set of axioms will fix anything even close to a unique interpretation for second order quantifiers governed by the full consequence relation.

Can the friend of full second-order logic do better? Perhaps. We can specify the full consequence relation model-theoretically. Call an inference from Δ to ϕ *approved* if and only if ϕ is a full model-theoretic consequence of Δ. Then we can say 'Let the second-order resources be interpreted however they must be in order to make all and only improved inferences valid.'

This response defuses the Intelligibility Argument. But it opens the door to skepticism. We *might* worry that the 'theory' (whether Henkin or second-order) isn't realizable at all. But it's plausible that any language which can be given a truth-conditional semantics can also be made meaningful, so we perhaps shouldn't worry to much about realizability. More worrisome is that, even if this gets us *an* interpretation for second-order quantification, it might get us an ontologically guilty one. For all that's been said, the only interpretation of '$\exists X^n \ldots$' that realizes the theory is 'there is a set-like entity X^n such that ...', and we're back in the arms of ontological guilt.

Of course, skeptical worries aren't arguments, and the friend of second-order quantification may simply take the ontological innocence of the interpretation on faith. At best we end in a dialectical stalemate: the friend of second-order logic certain she has the innocent interpretation she needs, and the foe certain she does not.

20.4 Set Theory in Sheep's Clothing

The arguments from ontological guilt aim to undercut the very idea of second-order quantification. The arguments in this section, by contrast, have no special beef with putting variables in predicate position. They aim instead to show that there is something objectionable about *full* second-order logic.

I mentioned in Sect. 20.2.2 that, if F is a system using the 'full' model theory, then there are sentences (CH) and (NCH) (using only logical vocabulary) such that, if the continuum hypothesis is true, $\vDash_F$ (CH), and if it is false, $\vDash_F$ (NCH). In other words, it treats either one or the other as a logical truth.

At its most basic, the 'Sheep's Clothing' worry is that no genuine consequence relation should do this. It comes in two forms. In one form, consequence relations shouldn't do this because it keeps them from Topic Neutrality. In the other, they shouldn't do this because it gets them into trouble with Normativity.

20.4.1 Topic Neutrality

Logical consequence is a relation between premises and conclusions; logical truth—or what I'll call *validity*—is a property of individual sentences. It's the property that a sentence has if and only if it's a logical consequence of any set of premises (including the empty set).

Every consequence relation has a corresponding validity property: ϕ is an *S-validity* if and only if it is an S-consequence of any premises whatsoever. A sentence ϕ is a *full second-order validity* iff $\Delta \vDash_F \phi$ for every set Δ. It is a *genuine validity* iff $\Delta \Rightarrow \phi$ for every Δ.

Topic Neutrality gave us a constraint on consequence relations, but it gives rise to a further constraint on validity: whether or not a given sentence is genuinely valid should not depend on the truths of any particular subject-matter. If full second-order logic is logic, then the full second-order validities are genuine validities. So if second-order logic is logic, whether a sentence is full-second-order valid shouldn't depend on the truths of any particular subject matter. But (goes the objection), it *does*—it depends on the truths of set theory, as the example of (CH) shows—so second-order logic isn't logic.

Put more precisely, this argument runs:

Topical Sheep's Clothing

(i) (CH) is a full second-order validity if and only if the continuum hypothesis is true.
(ii) If (i), then if second-order logic is logic, the logical validities depend on the truths of set theory.
(iii) The logical validities do not depend on the truths of set theory.
(iv) So full second-order logic is not logic.

Premise (i) is supported by a simple model-theoretic fact, and premise (iii) follows from the Topic Neutrality constraint. But what of premise (ii)?

Note first that, for premise (ii) to have any plausibility at all, the 'if and only if' in (i) needs to be stronger than a mere truth-functional one. It's not the mere observation that '(CH) is a full second-order validity' and 'the continuum hypothesis is true' have the same truth-value that underwrites (ii); rather, (i) is supposed to express some sort of fairly deep connection between the two.

The deep connection seems to be there: we can *prove* that, if (CH) is a second-order validity, the continuum hypothesis is true, and vice versa.[17] But that might not be enough to make second-order logic 'depend' on the truths of set theory in any objectionable way.

Why not? Well, there are similar connections between first-order logic and set theory, but this doesn't seem to make first-order logic unacceptably topic-sensitive. Consider, for instance, the first-order inference:

(18) $\forall x(Fx \rightarrow Gx)$
$\forall x(Gx \rightarrow Hx)$
$\therefore \forall x(Fx \rightarrow Hx)$

We can show that this inference is first-order valid if and only if the subset relation is transitive. This is because (on the standard model-theoretic semantics) $\ulcorner \forall x(\Pi x \rightarrow \Xi x) \urcorner$ is true if and only if the set of Π-satisfiers is a subset of the set of Ξ-satisfiers.[18] This connection does not seem to keep first-order logic from being logic, though, so it's not clear why the similar connection between (CH) and the continuum hypothesis is a problem for second-order logic.

Fans of the Topic Neutrality argument may object that the relation between (18) and transitivity is crucially different from that of (CH) and the continuum hypothesis. Although the biconditionals connecting consequence (or validity) to set theory are true in both cases, in the first case we rely on the validity of (18) to show that the subset relation is transitive, whereas in the second we rely on the (supposed) truth of the continuum hypothesis to show that (CH) is a full second-order validity. This suggests (goes the objection) that the dependencies are different in the two cases: the transitivity of subsethood depends on the validity of (18), whereas the validity of (CH) depends on the truth of the continuum hypothesis.

The observation about how we establish each biconditional seems right, as a mater of sociological fact. That's because we tend to use broadly first-order reasoning when reasoning model-theoretically about second-order logic.[19] It's not clear this sociological fact carries any weight, though. If we were happy to reason second-order, we might first formulate second-order set theory and then use the (assumed) validity of (CH) to show that the continuum hypothesis is true, and so

[17]Slightly more carefully, we can prove that, if $\vDash_F$ (CH), the continuum hypothesis is true, and vice versa; someone who follows the line of thought outlined in Sect. 20.4.2.3 will have room to resist concluding that if $\vDash_F$ (CH), (CH) is genuinely valid.

[18]Thanks to Aaron Cotnoir for suggesting this example to me.

[19]At least, if the 'we' are philosophers; Shapiro [21] argues at length that actual mathematical practice, which presumably includes that of model-theory, is rife with second-order reasoning. I cannot evaluate that claim here.

on. Whether we did this or not, we might think that the continuum hypothesis (or its negation) is the set-theoretic *result* of the (logically prior) truth of (CH) (or (NCH)). No one who takes this perspective should grant that (i) makes validities depend in any worrisome sense on set theory.

20.4.2 Normativity

An alternative argument sees troubles stemming from the interaction between full second-order logic's expressive power and the normativity of logic. According to the normativity constraint of Sect. 20.1.2, it's an error of reasoning to accept the premises of a genuinely valid argument while rejecting its conclusion. As a corollary, it's an error of reasoning to reject any genuine validities. If second-order logic is indeed logic, then either (CH) or (NCH) will be a genuine validity, and we will thus have logically-based epistemic obligations towards it. But the continuum hypothesis—and so, by extension, (CH) and (NCH)—seem radically epistemically unsettled in a way that doesn't sit well with these obligations.[20]

In argument form, these observations run:

Normative Sheep's Clothing

(i) If full second-order logic is logic, then either (CH) is a genuine validity or (NCH) is.
(ii) If (CH) is a genuine validity, then it's an error of reasoning to reject (CH).
(iii) If (NCH) is a genuine validity, then it's an error of reasoning to reject (NCH).
(iv) So if full second-order logic is logic, then it's either an error of reasoning to reject (CH) or an error in reasoning to reject (NCH).
(v) It is neither an error in reasoning to reject (CH) nor an error in reasoning to reject (NCH).
(vi) Therefore, full-second-order logic is not logic.

There are a number of routes for resistance; some look more promising than others. Let's go through a few in turn.

20.4.2.1 Agnosticism

We might be skeptical about premise (v). We might think that, given the deep epistemic openness about the continuum hypothesis, being opinionated on *either* count would be epistemically unwarranted. Agnosticism about the continuum hypothesis is plausibly the only epistemically responsible route. But agnosticism plausibly entails that we reject *neither* (CH) nor (NCH), so it is not compatible with premise (v).

[20] I assume that if it is an error in reasoning to A, then we have an epistemic obligation to not A. I will sometimes slide between error-talk and obligation-talk in the text.

Since premise (v) tempts us, those who would reject it should explain its appeal. The story, presumably, is that we get tempted because we conflate non-rejection with acceptance. If Normativity told us that we had to either accept (CH) or accept (NCH), that *would* be bad. But telling us to not reject a claim is far weaker than telling us to accept it, because the former but not the latter leaves agnosticism as an option.

I can imagine two further responses to this reply. The first response tries to beef up the original argument by appealing to a stronger form of normativity,

Strong Normativity: If $\Delta \Rightarrow \phi$, then anyone who accepts all of Δ (and meets some condition C) makes an error in reasoning if they do not accept ϕ, too,

and using this stronger form to argue that we must accept either (CH) or (NCH).

Why the rider 'and meets some condition C'? Because otherwise counterexamples will be too easy to come by. If the argument from Δ to ϕ is too long and complex, or if the logical structure of ϕ itself is too difficult for us to get into our heads, then plausibly we do nothing epistemically wrong if we don't accept it.

The need for the rider saddles the proponent of this response with the unenviable task of navigating between the cliffs of counterexample and the shoals of satisfiability. For if C is too weak, Strong Normativity will tell us we have obligations we clearly don't have. If C is too strong, though, then even if the beefed-up argument shows that some possible agents have to accept either (CH) or (NCH), if we mere mortals don't satisfy the condition, it doesn't show anything untoward about *our* epistemic obligations. Whether the responder can chart just the right course is something I'll not consider further here.

The second line of response is more subtle, and runs like this. 'Okay, so (v) is false. We should reject neither (CH) nor (NCH) because we lack the relevant evidence to come to an opinion. But this kind of impermissibility-of-rejection is *weaker* than the sort had by genuine validities. For instance, if I suppose that p and then reason myself into rejecting a genuine validity, that gives me epistemic warrant to reject the supposition. That's how *reductio* reasoning works. But supposing that p and then reasoning myself into rejecting something which, as a matter of fact, I don't have evidence to reject gives me no warrant for rejecting p. (CH) and (NCH) may be both impermissible-to-reject in the weaker, evidential sense, but not the stronger sense—but one of them would be if it were a genuine validity.'

Call a claim ϕ *reductio fodder* iff, if we reason to the rejection of ϕ under a supposition, that gives us warrant to reject the supposition. Then the second line of response suggests a modification of the original argument:

Reductio Sheep's Clothing

(i*) If full second-order logic is logic, then either (CH) is a genuine validity or (NCH) is.
(ii*) If (CH) is a genuine validity, then it is reductio fodder.
(iii*) If (NCH) is a genuine validity, then it is reductio fodder.
(iv*) So if full second-order logic is logic, then either (CH) is reductio fodder or (NCH) is.
(v*) Neither (CH) nor (NCH) is reductio fodder.
(vi*) Therefore, full-second-order logic is not logic.

It is far more difficult to deny (v*) than (v). Denying it would seem to undercut the considerably entrenched mathematical practice of showing various results to hold in the presence or absence of the continuum hypothesis. Presumably, if I suppose the truth of the continuum hypothesis, I can reason my way to the rejection of (NCH); but even if the continuum hypothesis is false (and so (NCH) reductio fodder), we would not think we could use reductio reasoning to come to know it this way.

I suspect many readers will find the original argument less compelling than this new one, so it so it might be prudent to focus on the latter from here on in. But I won't follow this wise counsel. The next few sections consider further objections to the original argument. This is of small moment, though, as the objections to be considered will apply, with only small modifications, to the Reductio argument too.

20.4.2.2 Indeterminacy

A number of philosophers think there's just no fact of the matter about whether the continuum hypothesis is true. Suppose that's right, and suppose second-order logic is logic. Then there will just be no fact of the matter as to whether (CH) or (NCH) is genuinely valid, either—even if, definitely, one of them is.

Normative Sheep's Clothing uses a form of dilemma reasoning, and many treatments of indeterminacy rule certain dillema-like inferences. Foes of the argument may be tempted to challenge the argument's validity on those grounds.

But that would be a mistake. Distinguish *direct* dilemma reasoning from the *indirect* sort, or 'proof by cases'. The direct form runs:

A or *B*.
If *A*, then *C*.
If *B*, then *C*.
Therefore, *C*.

The indirect form removes the second and third premises and replaces them with a procedure: suppose the antecedents and then prove the consequents under those suppositions. Proof by cases takes (direct) dilemma reasoning and replaces each conditional premise with an indirect proof.

Some treatments of indeterminacy invalidate indirect proof for conditionals. For instance, in supervaluational treatments we can derive ⌜Determinately, ϕ⌝ from ϕ, but we cannot then go on to prove ⌜$\phi \rightarrow$ Determinately, ϕ⌝. So indirect proof fails. In these systems, proof by cases fails for essentially the same reason. But direct disjunctive dilemma is valid in this system. Since Normative Sheep's Clothing relies on direct disjunctive dilemma, it can't be charged with straightforward invalidity, even on these treatments of indeterminacy.

If the continuum hypothesis is indeterminate, though, there may be a better criticism of Normative Sheep's Clothing in the neighborhood. Suppose that we are looking at a rose with a color in the penumbra between pink and red. Now consider this argument:

The Rose Argument

(i′) The rose is either red or pink.
(ii′) If the rose is red, we ought to believe that it is red.
(iii′) If the rose is pink, we ought to believe that it is pink.
(iv′) So we either ought to believe that the rose is red, or we ought to believe that it is pink.

Supposing vagueness is a species of indeterminacy, the Rose Argument parallels Normative Sheep's Clothing in interesting ways. The first premise is (on many theories of vagueness) true, and the argument valid. But the conclusion seems to commit us to the unacceptable claim that we ought to hold a definite opinion about whether a rose is red, even when it is a borderline case of being red.

One plausible diagnosis of the Rose Argument has it that premises (ii′) and (iii′) are false, but tempting. They are tempting because we tend to evaluate them by first imagining that we are in a position to assert 'the rose is red' or 'the rose is pink,' and then asking what we ought to believe in those cases. On many treatments of indeterminacy, we will only be in a position to assert either of these if the rose is *determinately* red or pink. We thus evaluate the conditionals by an imaginative analogue of indirect proof; but just as indirect proof gives us the wrong result when indeterminacy is involved, this imaginative process does, too.

In other words, we're tempted to accept the false premises (ii′) and (iii′) because our evaluative process confuses them with the true

(ii′) If the rose is determinately red, we ought to believe that it is red.
(iii′) If the rose is determinately pink, we ought to believe that it is pink.

These are true; but since it is *not* true that the rose is either determinately red or determinately pink, we can't use (ii″) and (iii″) to get (iv′).

If this is the right diagnosis of the Rose Argument, and if the continuum hypothesis is in fact indeterminate, then a similar diagnosis may fit Normative Sheep's Clothing.[21] For if the continuum hypothesis is indeterminate and second-order logic is logic, then although it may be true (and determinately true) that either (CH) or (NCH) is genuinely valid, it should be indeterminate which one is. But plausibly we only need to be governed by the determinate facts of logic: if it's unsettled which of several claims is genuinely valid, then our epistemic obligations ought not favor one over the other.

If *that's* right, though, then the Normativity Constraint from Sect. 20.1.2 needs to be tweaked: it's only an error in reasoning to accept all of Δ but reject ϕ when the latter is *determinately* a genuine consequence of the former. Since we can't get premises (ii) and (iii) of Normative Sheep's Clothing from this modified constraint, the argument fails.

[21] It may not be the right diagnosis, of course. Another plausible diagnosis [4] has it that the Rose Argument is sound and we should satisfy the obligations of (iv′) by getting ourselves into a position where it's indeterminate whether we believe that the rose is red or believe instead that the rose is pink.

20.4.2.3 Weakening the Logic

Premise (i) of the Normative Sheep's Clothing argument is underwritten by two claims. The first is that, if the continuum hypothesis is true, $\vDash_F$ (CH), and if it is false, $\vDash_F$ (NCH). The second is that, if full second-order logic is logic, then if $\vDash_F \phi$, ϕ is a genuine validity.

Motivation for the second claim stems from our definition of (full) second-order logic being *logic*: There is some second-order system SO where $\Delta \Rightarrow \phi$ iff $\Delta \vDash_{SO} \phi$. Motivation for the first claim comes from the following mathematical fact: First-order model theory, as standardly defined, is such that every model is a model of (CH) if the continuum hypothesis is true, and a model of (NCH) if the continuum hypothesis is false.

Here is a different conception of 'being logic', one endorsed by Field [5], who claims Kreisel [10] as inspiration. Recall that formal systems, as described in Sect. 20.1.1, include both a proof-theoretic and a model-theoretic component. On Field's picture, each of those components has an important job to do. When devising a formal system S for genuine logical consequence, we want to pick one where the proof system is 'genuinely sound': if $\Delta \vdash_S \phi$, then $\Delta \Rightarrow \phi$. And we want to choose a system with a model theory that is 'genuinely complete': if there is a model of Δ without ϕ—that is, if $\Delta \nvDash_S \phi$—then $\Delta \not\Rightarrow \phi$.

If the system is complete—if every model-theoretic consequence of the system is also one of its proof-theoretic consequences—then we get a 'squeezing argument' to show us that genuine consequence coincides with both the model-theoretic and the proof-theoretic relation. If the system is incomplete, though—as every full second-order system will be—the best we can say is that genuine consequence lies somewhere in between the proof- and model-theoretic relations. All proof-theoretic consequences are genuine consequences, and all genuine consequences are model-theoretic consequences. But that's all we can say.

From this perspective, we might think that full second-order logic is 'logic' iff there is a second-order system F which is genuinely sound and genuinely complete. In that case, saying that second-order logic is logic will only commit us to genuine consequence lying somewhere 'in between' $\vDash_F$ and $\vdash_F$. So, so long as F won't let us *derive* either (CH) or (NCH) (which, presumably, it won't), we can grant that, if the continuum hypothesis is true, $\vDash_F$ (CH), while insisting that it's still not a genuine validity—for this may be a case where the model-theoretic relation outstrips genuine consequence.

This undercuts the argument. But it does so at a cost. Friends of full second-order logic generally like it precisely because of its greater expressive power. Some may *want* either (CH) or (NCH) to be genuinely valid; they'll have no truck with this style of response. But even those who are happy to let (CH)'s or (NCH)'s validity go will still want full second-order logic to be genuinely stronger than the Henkin variety. If *all* we can say is that consequence lies between proof- and model-theory, then—since the proof theory can give us no logic stronger than a Henkin logic—we'll have no guarantee that genuine logic is any stronger than Henkin.

Of course, friends of this response may simply dig in their heels and insist that it *is* stronger. 'Formal systems are for helping us *investigate* consequence, not for *telling us what it is*. I don't guarantee anything about logic by specifying a formal system. Rather, the consequences are what they are; formal systems just help us reason about them. I happen to think that genuine consequence is almost-but-not-quite as strong as $\vDash_F$, and whether I can specify some model theory that exactly tracks it or not is beside the point.'

Fair enough; but we may legitimately wonder whether, after making this speech, the speaker can go on to learn very much at all from a formal system. Suppose he discovers that ϕ is a model-theoretic consequence of Δ. Can he conclude anything about their logical relationship? Unfortunately not—this may be one of those cases where model-theoretic consequence outstrips genuine consequence. In fact, his formal system will only let him be sure that ϕ is a consequence of Δ when ϕ is provable from (and then a Henkin consequence of) Δ. Of course, he might just insist he knows the genuine consequences when he sees them—but then he can forget the formal system entirely and just rely on his consequence-seeing faculties.

We may instead stick with our original gloss on 'being logic' and resist (i) by resisting the first claim that underwrote it. We cannot resist that claim by denying that, as standardly defined, every full second-order model is a model of (CH) if the continuum hypothesis is true and a model of (NCH) if the continuum hypothesis is false. That's a mathematical claim and not legitimately up for grabs. But we might try to deny it by giving a *non*-standard definition of a full model-theoretic consequence relation. We won't define $\vDash_F$ as 'truth in all second-order models', but as something more sophisticated cobbled together from these resources.

One way it might go stems from our thoughts about indeterminacy from Sect. 20.4.2.2. In that section, we considered the idea that the Normativity constraint ought to govern only determinate consequence. In other words, we granted that, determinately, if $\vDash_F \phi$, then ϕ is genuinely valid, but suggested that if it is indeterminate whether ϕ is valid, we are free to reject ϕ. But we could break the link earlier in the chain, insisting that if it is indeterminate whether $\vDash_F \phi$, then ϕ is *not* genuinely valid. We then identify the genuine validities as those which, determinately, are true on every model.

More precisely, given the standard model-theoretic relation $\vDash_F$, we can define another one, $\vDash_{F^D}$, where $\Delta \vDash_{F^D}$ iff, determinately, $\Delta \vDash_F$. The formal system F^D is the one just like F except that we swap $\vDash_{F^D}$ for $\vDash_F$. Clearly, F^D should count as second-order system if F does.[22] The friend of second-order logic can think that it 'is logic' because $\Delta \vDash_{F^D} \phi$ iff $\Delta \Rightarrow \phi$ without being committed to either (CH) or (NCH) being genuinely valid.

Unlike the Kreisel-based rejection of (i), this move gives us a clearer picture of exactly which model-theoertic consequences correspond to genuine consequences:

[22]Should it count as a *full* second-order system? It's genuinely weaker than the (usual) full system, but stronger than Henkin systems (see the next note). I doubt usage is fixed enough to settle this question.

they are the *determinate* ones. Even if only the determinate holding of $\vDash_F$ gives us genuine consequences, that's enough to show that genuine second-order consequence outstrips Henkin consequence.[23] And, insofar as we have a fairly good grip on which model-theoretic results depend on claims of dubious determinacy and which do not, we can use the second-order model theory as a tool to investigate genuine consequence.

20.4.2.4 Rejecting Normativity

A final option is nuclear: simply reject the normativity constraint.

Let's be a bit more precise. Aside from a brief parenthetical remark in Sect. 20.1.2, I have been talking as though there is a unique genuine consequence relation that we care about. But that's not at all obvious. There may be a number of different relations, each of which deserves to be called a consequence relation.

If that's right, it may be that not all such relations obey, or obey equally well, all of the constraints of Sect. 20.4.2.2. Perhaps, for instance, the kind of normativity we care about in logic ones apart from the kind of necessity we care about. This idea isn't new: David Kaplan ([9]; see also [20]) suggests that 'I am here now' is a kind of logical validity, and the reason is at least partly because no one should reject it. Yet it's clearly not necessary: I may be here now, but I could have been somewhere else instead.

If normativity and necessity come apart, then there may be several candidate 'logical consequence' relations: one that tracks the distinctive normativity of logic, one that tracks the distinctive necessity of logic, and so on. Shapiro [22, pp. 772–773] has suggested something more-or-less along these lines. And he has suggested that full second-order logic is logic precisely because it tracks the distinctive necessity—normativity be hanged.[24]

20.5 Scorecard

Rather than trying to exonerate second-order logic, I have merely provided a brief for the defense. That defense will find some challenges more worrying than others. In particular, the Intelligibility and Normative Sheep's Clothing arguments present

[23] The quick-and-dirty way to show this is to note that full second-order logic has sentences that characterize infinite models. So long as our notion of infinity isn't itself indeterminate, this means that (determinately) we have a sentence that is entailed by an infinite set ('there is at least one thing,' 'there is at least two things,' ...) but not any of its finite subsets. So, unlike Henkin systems, the system F^D is not compact.

[24] This may not be an entirely fair characterization of Shapiro's view, in large part because he is considering a normative constraint somewhat stronger than the one outlined in Sect. 20.1.2. Still, this captures the basic idea, and is a move available here to friends of second-order logic worried about our weaker Normativity constraint.

deeper difficulties than Textbook Quineanism or Topical Sheep's Clothing. Still, the friend of second-order logic has several defensive avenues available. I leave it to her to decide which, if any, she wishes to take.

References

1. Barcan Marcus, R. 1972. Quantification and ontology. *Nous* 6(3): 240–250.
2. Beall, J.C., and G. Restall. 2006. *Logical pluralism*. Oxford: Oxford University Press.
3. Boolos, G. 1975. On second-order logic. *The Journal of Philosophy* 72(16): 509–527.
4. Dorr, C. 2003. Vagueness without ignorance. *Philosophical Perspectives* 17(1): 83–113.
5. Field, H. 1991. Metalogic and modality. *Philosophical Studies* 62(1): 1–22.
6. Field, H. 2008. *Saving truth from paradox*. Oxford: Oxford University Press.
7. Haack, S. 1978. *Philosophy of logics*. Cambridge: Cambridge University Press.
8. Henkin, L. 1950. Completeness in the theory of types. *The Journal of Symbolic Logic* 15(2): 81–91.
9. Kaplan, D. 1989. Demonstratives. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein, 481–563. New York: Oxford University Press.
10. Kreisel, G. 1967. Informal rigor and completeness proofs. In *Problems in the philosophy of mathematics*, ed. I. Lakatos, 138–171. Amsterdam: North-Holland.
11. Lewis, D. 1970. How to define theoretical terms. *The Journal of Philosophy* 67: 427–446. Reprinted in Lewis, D. 1983. *Philosophical papers*, vol. 1, 78–95. Oxford: Oxford University Press.
12. Prior, A.N. 1971. *Objects of thought*. Oxford: Clarendon.
13. Quine, W. 1940. *Mathematical logic*. Cambridge, MA: Harvard University Press.
14. Quine, W. 1960. *Word and object*. Cambridge: MIT.
15. Quine, W. 1970. *Philosophy of logic*. Englewood Cliffs: Prentice-Hall.
16. Rayo, A. (2008). On specifying truth-conditions. *The Philosophical Review* 117: 385–443.
17. Rayo, A., and T. Williamson. 2003. A completeness theorem for unrestricted first-order languages. In *Liars and heaps: New essays on paradox*, ed. J.C. Beall, chapter 15, 331–356. Oxford: Oxford University Press.
18. Rayo, A., and S. Yablo. 2002. Nominalism through de-nominalization. *Noûs* 35(1): 74–92.
19. Rossberg, M. Forthcoming. Somehow things do not relate: On the interpretation of polyadic second-order logic. *The Journal of Philosophical Logic*.
20. Russell, G. 2008. One true logic? *The Journal of Philosophical Logic* 37(8): 593–611.
21. Shapiro, S. 1991. Foundations without foundationalism. Oxford: Oxford University Press.
22. Shapiro, S. 2005. Higher-order logic. In *The Oxford handbook of philosophy of mathematics and logic*, ed. S. Shapiro, 751–780. Oxford: Oxford University Press.
23. van Inwagen, P. 1981. Why I don't understand substitutional quantification. *Philosophical Studies* 39: 281–285. Reprinted in van Inwagen, P. 2001. *Ontololgy, identity, and modality*, 32–36. Cambridge: Cambridge University Press.
24. van Inwagen, P. 1998. Meta-ontology. *Erkenntnis* 38: 223–250. Reprinted in van Inwagen, P. 2001. *Ontololgy, identity, and modality*, 13–31. Cambridge: Cambridge University Press.
25. van Inwagen, P. 2004. A theory of properties. In *Oxford studies in metaphysics*, vol. 1, ed. D.W. Zimmerman, 107–138. Oxford: Oxford University Press.
26. Williams, J. 2010. Fundamental and derivative truths. *Mind* 119(473): 103–141.

Chapter 21
∀ and ω

Elia Zardini

Abstract I first briefly rehearse the two substructural solutions that I've elsewhere proposed to the semantic and vagueness paradoxes. I then ask what the correct principle of universal generalisation is. The traditional answer to this question is represented by the familiar principle to the effect that, provided that τ does not occur free in either Γ, Δ or φ, if $\Gamma \vdash \Delta, \varphi_{\tau/\xi}$ holds, $\Gamma \vdash \Delta, \forall\xi\varphi$ holds. I argue for interpreting such principle as in effect licencing the inference from 'anything' to 'everything'. I then proceed to offer five arguments against that inference. The first three arguments rely on considerations concerning the preface paradox, the failure of agglomeration for counterfactual implication and free-choice permission respectively. The last two arguments connect back with the semantic and vagueness paradoxes. I show how the inference from 'anything' to 'everything' would wreak havoc for the workings both of my non-contractive solution to the semantic paradoxes and of my non-transitive solution to the vagueness paradoxes. I then inquire into what a more adequate generalisation principle should be, and argue in favour of a suitably generalised version of the ω-rule, defending it from several prominent objections. I then trace back the quantificational phenomena studied in the paper, in particular those most directly related to the semantic and vagueness paradoxes, to their sentential root concerning the behaviour of conjunction. I sketch a metaphysical view making sense of the failure of the conjunctive analogue of the traditional generalisation principle, and close by bringing out some positive implications such view has for our logical freedom.

E. Zardini (✉)
FCT Research Fellow, LanCog, Language, Mind and Cognition Research Group, Centro de filosofia, Universidade de Lisboa, Lisbon, Portugal
e-mail: elia.zardini@campus.ul.pt

© Springer International Publishing Switzerland 2015
A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library 373,
DOI 10.1007/978-3-319-18362-6_21

21.1 Semantic Paradoxes and Failure of Contraction

In [26–28, 32, 35, 36, 38], I've developed a theory solving the *semantic paradoxes*. The theory, **LW**, basically consists of the *multiplicative sentential fragment of affine logic* supplemented by the metarules for *universal quantification*:

$$\frac{\Gamma, \varphi_{\upsilon'/\xi}, \varphi_{\upsilon''/\xi}, \varphi_{\upsilon'''/\xi} \ldots \vdash_{\mathbf{LW}} \Delta}{\Gamma, \forall\xi\varphi \vdash_{\mathbf{LW}} \Delta} \ \forall\text{-L}$$

$$\frac{\Gamma' \vdash_{\mathbf{LW}} \Delta', \varphi_{\upsilon'/\xi} \qquad \Gamma'' \vdash_{\mathbf{LW}} \Delta'', \varphi_{\upsilon''/\xi} \qquad \Gamma''' \vdash_{\mathbf{LW}} \Delta''', \varphi_{\upsilon'''/\xi} \ldots}{\Gamma', \Gamma'', \Gamma''' \ldots \vdash_{\mathbf{LW}} \Delta', \Delta'', \Delta''' \ldots \forall\xi\varphi} \ \forall\text{-R}$$

(where, roughly, φ_{τ_0/τ_1} is the result of replacing in φ all free occurrences of τ_1 with free occurrences of τ_0 and '$\upsilon', \upsilon'', \upsilon''' \ldots$' and its like denote a canonical enumeration of the totality of linguistic entities of a certain kind)[1,2] and by the metarules for *truth*:

$$\frac{\Gamma, \varphi \vdash_{\mathbf{LW}} \Delta}{\Gamma, V\ulcorner\varphi\urcorner \vdash_{\mathbf{LW}} \Delta} \ V\text{-L} \qquad\qquad \frac{\Gamma \vdash_{\mathbf{LW}} \Delta, \varphi}{\Gamma \vdash_{\mathbf{LW}} \Delta, V\ulcorner\varphi\urcorner} \ V\text{-R}$$

(where V is a truth predicate and $\ulcorner\varphi\urcorner$ a canonical object-language singular term denoting φ).

LW is so-called because it lacks the metarules of *contraction*:

$$\frac{\varphi, \varphi \vdash \psi}{\varphi \vdash \psi} \ \text{W-L} \qquad\qquad \frac{\varphi \vdash \psi, \psi}{\varphi \vdash \psi} \ \text{W-R}$$

For example, failure of contraction blocks a standard *Liar* paradox as follows. Where λ is $\neg V\ulcorner\lambda\urcorner$, $\lambda \vdash_{\mathbf{LW}} \lambda$ holds and $\lambda \vdash_{\mathbf{LW}} \neg\lambda$ holds, and so $\lambda, \lambda \vdash_{\mathbf{LW}} \lambda \,\&\, \neg\lambda$ holds, but, because of failure of contraction, $\lambda \vdash_{\mathbf{LW}} \lambda \,\&\, \neg\lambda$ does not hold, and the familiar paradoxical reasoning thus breaks down.

Restricting contraction allows **LW** to have (among other things) *transparent* truth (that is, φ is fully intersubstitutable with $V\ulcorner\varphi\urcorner$), *Boolean* negation (that is, $\neg\varphi$ is genuinely exclusive and exhaustive with respect to φ), the *full deduction theorem* for (material) implication (that is, $\Gamma, \varphi \vdash \Delta, \psi$ holds iff $\Gamma \vdash \Delta, \varphi \supset \psi$ holds) and *extensional* theories of conjunction, disjunction and quantification (that is, the holding of both φ and ψ is necessary and sufficient for the holding of $\varphi \,\&\, \psi$, the

[1] Obviously, these metarules for universal quantification only make sense under certain assumptions about the contextually relevant domain of discourse and the language, which I'll make explicit and discuss in Sect. 21.5.

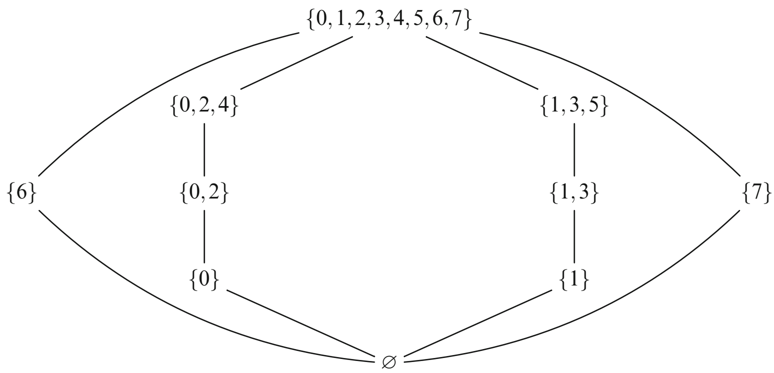
[2] Throughout, I assume that *particular quantification* is defined in the familar way using universal quantification and negation. Although I'll save the reader the gory details, essentially everything I'll say about universal quantification has an analogue concerning particular quantification.

holding of either φ or ψ is necessary and sufficient for the holding of $\varphi \vee \psi$, the holding of the totality of $\varphi_{\upsilon'/\xi}, \varphi_{\upsilon''/\xi}, \varphi_{\upsilon'''/\xi} \ldots$ is necessary and sufficient for the holding of $\forall\xi\varphi$). No other theory I know of has all of these.

21.2 Vagueness Paradoxes and Failure of Transitivity

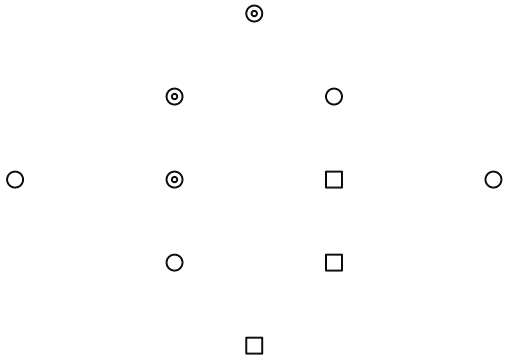
In [23–25, 29, 31, 34, 37], I've developed a theory solving the *vagueness paradoxes*. The theory is best thought of as a *non-logical theory*, $\mathcal{N}$, plus a *logic*, **LS**. $\mathcal{N}$ basically consists, for example, of $B0 \mathbin{\&} \neg B999, 999 \mathbin{\&} \forall x(Bx \supset Bx')$ (read 'Bi' as 'A man with i hairs is bald' and make the customary simplifying assumption that baldness is only a matter of number of hairs on one's scalp). **LS** is less familiar, and can be constructed as follows. An **LS**-model $\mathfrak{M}$ is a 7ple $\langle U_{\mathfrak{M}}, V_{\mathfrak{M}}, \preceq_{\mathfrak{M}}, D_{\mathfrak{M}}, T_{\mathfrak{M}}, \mathbf{neg}_{\mathfrak{M}}, \mathbf{int}_{\mathfrak{M}}\rangle$, where:

- $U_{\mathfrak{M}}$ is a *domain* of objects;
- $V_{\mathfrak{M}}$ is a set of *values* representable as: $\{X : X \in \mathbf{pow}(\{i : 0 \leq i \leq 7\})$ and, if $X \neq \{i : 0 \leq i \leq 7\}$, either, [[for every $i \in X$, i is even][3] and, [for every i and j, if $i \in X$ and ≤ 4 and j is even and $< i$, $j \in X$] and, [for every i and $j \in X$, $|i-j| < 6$]] or, [[for every $i \in X$, i is odd] and, [for every i and j, if $i \in X$ and ≤ 5 and j is odd and $< i$, $j \in X$] and, [for every i and $j \in X$, $|i-j| < 6$]]$\}$;
- $\preceq_{\mathfrak{M}}$ is a *partial order* on $V_{\mathfrak{M}}$ representable as: $\{\langle X, Y\rangle : X \subseteq Y\}$. Thus, $V_{\mathfrak{M}}$ and $\preceq_{\mathfrak{M}}$ jointly constitute the lattice depicted by the following Hasse diagram:

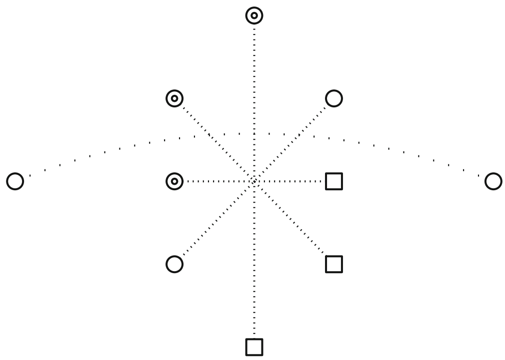


- $D_{\mathfrak{M}}$ is a set of *designated* values while $T_{\mathfrak{M}}$ is a set of *tolerated* values, with $D_{\mathfrak{M}} \subset T_{\mathfrak{M}}$. Indicating designated values with doubly circular nodes, tolerated but not designated values with simply circular nodes and neither designated nor tolerated values with square nodes, they can be depicted as:

[3]Throughout, I use square brackets to disambiguate constituent structure.



- **neg**$_{\mathfrak{M}}$ is a *negation* operation on $V_{\mathfrak{M}}$. Indicating it with pointed edges, it can be depicted as:



- **int**$_{\mathfrak{M}}$ is an *interpretation* function from the union of the Cartesian products of the sets of *i*ary atomic predicates with the sets of *i*tuples of members of $U_{\mathfrak{M}}$ to $V_{\mathfrak{M}}$.

int$_{\mathfrak{M}}$ can be extended to a full *valuation* function **val**$_{\mathfrak{M}}$ (relative to assignments) in the usual way (using glb in $\preceq_{\mathfrak{M}}$ for interpreting conjunction and universal quantification and using **neg**$_{\mathfrak{M}}$ for interpreting negation).[4] $\Gamma \vdash_{\mathbf{LS}} \Delta$ holds iff, for every **LS**-model $\mathfrak{M}$ and assignment **ass**, if, for every $\varphi \in \Gamma$, $\mathbf{val}_{\mathfrak{M},\mathbf{ass}}(\varphi) \in D_{\mathfrak{M}}$, then, for some $\psi \in \Delta$, $\mathbf{val}_{\mathfrak{M},\mathbf{ass}}(\psi) \in T_{\mathfrak{M}}$.

[4]In addition to particular quantification (see Footnote 2), I assume that disjunction and implication are defined in **LS** in the familar way using conjunction and negation. Notice that such definitions are also available in the case of **LW**: in both **LW** and **LS**, *the usual logical operations can essentially be reduced to conjoining and negating* (such reduction does encompass also the usual logical operation of universal quantification, which, in a more informal sense that should become clear as this paper progresses, can be reduced to a kind of infinite conjunction).

LS is so-called because it lacks the metarule of *transitivity*[5]:

$$\frac{\varphi \vdash \psi \qquad \psi \vdash \chi}{\varphi \vdash \chi}\,\text{S}$$

For example, failure of transitivity blocks a standard *Sorites* paradox as follows. Where β_i is *Bi*, $\mathcal{N} \vdash_{\mathbf{LS}} \mathcal{N} \,\&\, \beta_1$ holds and $\mathcal{N} \,\&\, \beta_1 \vdash_{\mathbf{LS}} \beta_2$ holds, but, because of failure of transitivity, $\mathcal{N} \vdash_{\mathbf{LS}} \beta_2$ does not hold, and the familiar paradoxical reasoning thus breaks down.

Restricting transitivity allows $\mathcal{N}$ with **LS** ($\mathcal{N}_{\mathbf{LS}}$) to have (among other things) *tolerant* baldness (that is, $\mathcal{N}$ and β_i entail β_{i+1}), *Boolean* negation, the *full deduction theorem* for implication and *extensional* theories of conjunction, disjunction and quantification. No other theory I know of has all of these.

Different as they may be in many other respects, as I'll explain **LW** and $\mathcal{N}_{\mathbf{LS}}$ agree in rejecting the traditional principle of *universal generalisation*. After going through some preliminaries in Sect. 21.3, I'll argue in Sect. 21.4 that the traditional principle is problematic for several independent reasons, and, in Sect. 21.5, I'll propose and defend a well-known alternative that also happens to be the one available for **LW** and $\mathcal{N}_{\mathbf{LS}}$. I'll close in Sect. 21.6 by offering an explanation of why, at least from a logical and metaphysical point of view shared by **LW** and $\mathcal{N}_{\mathbf{LS}}$, the traditional principle fails.

21.3 Universal Generalisation

It is no mystery what the *truth conditions* of a sentence $\forall\xi\varphi$ are: roughly, such a sentence is true iff the *totality* of objects in the contextually relevant domain of discourse satisfy φ. Henceforth assuming that the *basic principles of an adequate logic* for 'everything' should consist of an *instantiation* principle and a *generalisation* principle, it is usually thought that it is equally no mystery what such principles are. The traditional candidate for being the instantiation principle is the *from-everything-to-a-certain-thing* rule:

($\text{E}^{\Rightarrow}\text{C}$) If τ is a contextually relevant singular term of the language, $\forall\xi\varphi \vdash \varphi_{\tau/\xi}$ holds,

[5]Before my works on vagueness referenced above in the text, Alan Weir already developed a broadly non-transitive theory of truth (and sets) with possible application to vagueness (see [19] for a recent presentation). A crucial difference between Weir's theory and mine is that *Weir's still validates* S *and only rejects its strengthening with side premises and conclusions*. Without going into its merits as a theory of truth (and sets), I think that the validity of S does make the possible application of Weir's theory to vagueness very problematic: for one thing, it prevents it from endorsing the straightforward solution to the Sorites paradox that I'll mention below in the text.

and the traditional candidate for being the generalisation principle is the *from-a-certain-thing-without-assumptions-to-everything* right-metarule:

($^{\Rightarrow}$C$_{\text{A}}$E) If τ does not occur free in either Γ, Δ or φ, if $\Gamma \vdash \Delta, \varphi_{\tau/\xi}$ holds, $\Gamma \vdash \Delta, \forall\xi\varphi$ holds.

Let's focus for the time being on ($^{\Rightarrow}$C$_{\text{A}}$E). Notice that 'τ' as it occurs in ($^{\Rightarrow}$C$_{\text{A}}$E) is sometimes understood, more naturally, as ranging over *variables*, while it is some other times understood, more parsimoniously, as ranging over any contextually relevant *atomic singular term* of the language. While both understandings are technically viable, and while possible differences between them will not be relevant for our discussion, to fix ideas I'll henceforth presuppose the first understanding of ($^{\Rightarrow}$C$_{\text{A}}$E). This obviously does not yet settle the issue of how to understand such occurrences of τ (in jargon known as occurrences of τ as an "*eigenvariable*" or "*parametre*"), an issue to which we now turn.

($^{\Rightarrow}$C$_{\text{A}}$E) is typically glossed along the following lines: if from certain assumptions one can derive that *x* is *F without making any specific assumption about x*, then from the same assumptions one can derive that everything is *F*. We can then sharpen our issue as the issue of how to understand '*x* is *F*' in a context in which no specific assumption is made about *x*. At this point, a natural and common move—whose correctness I'll henceforth assume—is to understand '*x* is *F*' in such a context as being tantamount to 'An *arbitrary* thing is *F*'. In turn, 'an arbitrary thing' has notoriously given rise to all sorts of speculations (see Sect. 21.5 for an example), but an attractive conception—whose correctness again I'll henceforth assume—appeals to a notion deeply entrenched in *ordinary* and *scientific* (especially, *mathematical*) thought, if rarely singled out for *logical* investigation, understanding 'An arbitrary thing is *F*' as being tantamount to '*Anything* is *F*'.

On this scheme, the apparently *referential* construction 'An arbitrary thing is *F*' is understood in terms of—and its apparent referentiality is explained away by—the more fundamental, *non-referential* and instead *quantificational* construction 'Anything is *F*', which we can formally write as $\mathcal{A}\xi\varphi$ and understand as governed by the *from-anything-to-a-certain-thing* rule:

(A$^{\Rightarrow}$C) If τ is a contextually relevant singular term of the language, $\mathcal{A}\xi\varphi \vdash \varphi_{\tau/\xi}$ holds

and the *from-a-certain-thing-without-assumptions-to-anything* right-metarule:

($^{\Rightarrow}$C$_{\text{A}}$A) If τ does not occur free in either Γ, Δ or φ, if $\Gamma \vdash \Delta, \varphi_{\tau/\xi}$ holds, $\Gamma \vdash \Delta, \mathcal{A}\xi\varphi$ holds.[6]

[6]It might be worried that the proposed quantificational understanding of eigenvariables spoils much of the point of ($^{\Rightarrow}$C$_{\text{A}}$E) and ($^{\Rightarrow}$C$_{\text{A}}$A), which would consist in making available the full power of *sentential* (i.e. non-quantificational) reasoning in manipulating a sentence containing an

Assuming that, no matter which singular terms occur free in φ, we can always enrich the language with a new contextually relevant singular term that does not so occur, ($A^{\Rightarrow}C$) and ($^{\Rightarrow}C_AE$) entail the *from-anything-to-everything* rule:

($A^{\Rightarrow}E$) $\mathcal{A}\xi\varphi \vdash \forall\xi\varphi$ holds.

---

eigenvariable in preparation for a generalisation. Focussing on ($^{\Rightarrow}C_AA$), it might be worried, for example, that, while an alleged success story of modern logic is to justify $\mathcal{A}\xi\Phi\xi \vdash \mathcal{A}\xi(\Phi\xi \vee \Psi\xi)$ through the usual eigenvariable reasoning involving the inference from $\Phi\tau$ to $\Phi\tau \vee \Psi\tau$, which allegedly can in turn be understood following the standard, non-quantificational semantics for sentential logic, the proposed quantificational understanding of eigenvariables would make a hash of any such justification, since it would understand the inference from $\Phi\tau$ to $\Phi\tau \vee \Psi\tau$ as tantamount to $\mathcal{A}\xi\Phi\xi \vdash \mathcal{A}\xi(\Phi\xi \vee \Psi\xi)$ itself. I think there are two serious problems with this description of the situation. Firstly, a sentence containing an eigenvariable is naturally understood so that the eigenvariable always takes *wide scope* (so much has already been implicitly assumed by the "typical gloss" in the last paragraph in the text), which clashes with the standard semantics for sentential logic, for example implying that, if τ occurs as an eigenvariable in $\varphi_{\tau/\xi}$ and $\neg\varphi_{\tau/\xi}$, these are not guaranteed to be contradictories (this is borne out by "the natural and common move" in the last paragraph in the text, since 'An arbitrary thing is *F*' and 'An arbitrary thing is not *F*' can both be false). Zooming in on our example, if τ occurs as an eigenvariable in $\Phi\tau \vee \Psi\tau$, that is naturally and commonly read not as the disjunction 'Either an arbitrary thing is Φ or an arbitrary thing is Ψ' (the only kind of reading envisaged by the standard semantics for sentential logic), but as the complex predication 'An arbitrary thing either is Φ or is Ψ'. Secondly, it is trivially if perhaps surprisingly not the case that a sentence containing an eigenvariable is subject to the *full* power of sentential reasoning: for example, it cannot be assumed as such (i.e. as containing an eigenvariable). "Eigenvariable logic" is deeply deviant. Zooming in again on our example, this raises a difficulty for the justification of $\mathcal{A}\xi\Phi\xi \vdash \mathcal{A}\xi(\Phi\xi \vee \Psi\xi)$ through the usual eigenvariable reasoning, since, on its most natural version, such reasoning requires chaining together $\mathcal{A}\xi\Phi\xi \vdash \Phi\tau$ and $\Phi\tau \vdash \Phi\tau \vee \Psi\tau$, which would seem to be in danger of equivocating on τ, given that this occurs as an eigenvariable in the former argument but not in the latter. The situation might start to look unexpectedly dire, but I think it improves considerably if we reflect that *sentential logic can non-standardly be interpreted as a quantificational logic of sorts*, to the effect that, roughly, '$\Gamma, \varphi_0, \varphi_1, \varphi_2 \ldots \vdash \psi(\tau_0, \tau_1, \tau_2 \ldots)$ holds' (where $\tau_0, \tau_1, \tau_2 \ldots$ are the free singular terms that are target of the quantificational understanding and $\varphi_0, \varphi_1, \varphi_2 \ldots$ are the assumptions, if any, in which such terms occur free) means that 'Given Γ, any objects τ_0, τ_1, τ_2 such that $\varphi_0, \varphi_1, \varphi_2 \ldots$ are such that ψ' is a logical truth. Such interpretation has, among other things, the virtue of providing *a uniform understanding of free variables*: τ can always be understood as 'any object τ', the only difference when it occurs as an eigenvariable being that no 'such that'-restrictions are placed on it: *discharging assumptions* is interpreted as *lifting restrictions*. (Notice that the interpretation would remain uniform for the relevant variables at work in the dual of ($^{\Rightarrow}C_AE$) for particular quantification.) I propose that it is this non-standard interpretation of "sentential" logic that is operative when we use it to manipulate sentences containing eigenvariables. The interpretation straightforwardly accounts for the two facts noted above. (It is at least perhaps worth recording that, for example, a sentence containing an eigenvariable cannot be assumed as such because this would change its interpretation as conclusion by adding a restriction—$\varphi_{\tau/\xi} \vdash \varphi_{\tau/\xi}$ does hold, but it means something along the lines of the claim that the restricted quantification 'Any object τ such that $\varphi_{\tau/\xi}$ is such that $\varphi_{\tau/\xi}$' is a logical truth). Zooming in for the last time on our example, the interpretation provides a satisfactory analysis of the justification of $\mathcal{A}\xi\Phi\xi \vdash \mathcal{A}\xi(\Phi\xi \vee \Psi\xi)$ through the usual eigenvariable reasoning. One fundamental element of such justification is $\Phi\tau \vdash \Phi\tau \vee \Psi\tau$, which, contrary to what the worry under discussion would have, is now interpreted as the claim that 'Any object τ such that $\Phi\tau$ is such that $\Phi\tau \vee \Psi\tau$' is a logical truth (you may gloss this element as the fact that the property of being *F* entails the property of either being *F* or being *G*). The other

Moreover, assuming that, no matter which singular terms occur free in either Γ, Δ or φ, we can always enrich the language with a new contextually relevant singular term that does not so occur, by an uncontroversial (that is, **LS**-acceptable) application of transitivity of logical consequence ($A^{\Rightarrow}C$) and ($^{\Rightarrow}C_AE$) entail the *from-anything-to-everything* right-metarule:

($^{\Rightarrow}$AE) If $\Gamma \vdash \Delta, \mathcal{A}\xi\varphi$ holds, $\Gamma \vdash \Delta, \forall\xi\varphi$ holds.[7]

21.4 The Difference between Anythingness and Everythingness

Trying to make sure to hit the target, I'm going to give no less than five arguments to the effect that there is a difference between anythingness and everythingness,[8] to the extent that ($^{\Rightarrow}C_AE$) fails. As we've seen in Sect. 21.3, under uncontroversial assumptions ($^{\Rightarrow}C_AE$) entails ($A^{\Rightarrow}E$) and ($^{\Rightarrow}$AE): the first four arguments target nothing less than ($A^{\Rightarrow}E$), and so establish nothing less than the stronger claim that anythingness does not even entail everythingness (all this with the proviso entered in Footnote 14); the last argument only targets ($^{\Rightarrow}$AE), and so only establishes the weaker claim that there is a difference between anythingness and everythingness.[9,10]

The first three arguments to the effect that there is a difference between anythingness and everythingness draw on considerations that are relatively independent of the semantic and vagueness paradoxes as well as of the theories solving them

fundamental element of such justification is a suitable version of *Barbara* (*ubi est logica moderna victoria tua?*), which allows the inference from the logical truth of 'Given that anything is Φ, any object τ is such that $\Phi\tau$' (which is how $\mathcal{A}\xi\Phi\xi \vdash \Phi\tau$ is now interpreted) and the logical truth of 'Any object τ such that $\Phi\tau$ is such that $\Phi\tau \vee \Psi\tau$' to the logical truth of 'Given that anything is Φ, any object τ is such that $\Phi\tau \vee \Psi\tau$' (you may gloss this element as the fact that 'anything' is upwards monotonic). Thanks to Sven Rosenkranz for discussion of this worry.

[7] Alternatively, without appeal to ($A^{\Rightarrow}C$) or transitivity of logical consequence, we can observe that, given our assumptions about the relation between 'an arbitrary thing' and 'anything', the *from-anything-to-a-certain-thing-without-assumptions* right-metarule:

($^{\Rightarrow}AC_A$) If τ does not occur free in either Γ, Δ or φ, if $\Gamma \vdash \Delta, \mathcal{A}\xi\varphi$ holds, $\Gamma \vdash \Delta, \varphi_{\tau/\xi}$ holds

is equally compelling. Assuming that, no matter which singular terms occur free in either Γ, Δ or φ, we can always enrich the language with a new singular term that does not so occur, ($^{\Rightarrow}AC_A$) and ($^{\Rightarrow}C_AE$) entail ($^{\Rightarrow}$AE).

[8] I'll henceforth unabashedly follow the venerable tradition of hypostasising logical concepts (see e.g. [9]).

[9] The difference, but none of these arguments, is anticipated by [16], chapter 5.

[10] I stress once and for all that the point I'm going to make is that there is a certain difference in *non-linguistic* reality. I think that English 'anything' and 'everything' track pretty well such difference (and so, throughout, I use them to mark it), and that certain facts about natural languages might (as usual) be relevant to the issue, but the point is not about English 'anything' and 'everything' (nor about any other linguistic expression of any other language).

that I've advocated in Sects. 21.1 and 21.2 respectively. Firstly, consider the familiar point from discussion of the *preface paradox* (see [12], p. 205) to the effect that, for example, one can justifiedly [believe of every belief one has that it is true without believing that every belief one has is true], even if one knows which beliefs one has (a point which I'll henceforth assume without further argument to be correct). The point is familiar and yet still widely controversial. I think that at least one source of the controversiality consists in the fact that it admittedly seems to border on the *incoherent* for one to have the belief that belief #0 is true, the belief that belief #1 is true, the belief that belief #2 is true ... without having any *general* belief about the truth of one's beliefs[11]—if one believed all those things about *particular* beliefs but believed no more, one would be *objectionably blind to a pattern* that is in fact pinned down by what one believes (and that however transcends any one of those particular beliefs).[12] But what could that general belief be if not the belief that every belief one has is true? Here is another way to put the worry. Presumably, when one *coherently* believes of every belief one has that it is true, one is thereby in an *individual, distinctive state of mind* that goes beyond the state of mind constituted by the belief that belief #0 is true, beyond the state of mind constituted by the belief that belief #1 is true, beyond the state of mind constituted by the belief that belief #2 is true...—the relevant portion of one's cognitive life does *not reduce to a schizofrenic plurality of states of mind each only about a particular belief*. Moreover, the state of mind one is in is *representational*: it represents a way in which the world could be. But what, in the example, could that way be if not the way the world is iff every belief one has is true?

This worry admits of a natural resolution that manages to be faithful to its presuppositions while preserving the point forcefully brought out by the preface paradox: the general belief one has is not the belief that *every* belief one has is true, but the belief that *any* belief one has is true, and the way one's state of mind represents the world to be is not the way the world is iff *every* belief one has is true, but the way the world is iff *any* belief one has is true. However, if (A$\Rightarrow$E)

[11]I think that the most natural understanding of the example (which has the virtue of avoiding distracting issues of *ungroundedness*) is one in which the belief that belief #0 is true, the belief that belief #1 is true, the belief that belief #2 is true ... and any related general belief are themselves *not* ones of belief #0, belief #1, belief #2... I'll henceforth tacitly assume this distinction among beliefs, leaving context to disambiguate what exactly is covered by a certain occurrence of 'belief' or of its relatives.

[12]Some may be tempted to require further that the general belief actually be the *ground* on which the beliefs about particular beliefs *depend*. But, less importantly, that would seem to turn things upside down in epistemic structure: as far as I can tell, there is no legitimate *default presumption* about the general truth of one's beliefs, but only an extremely compelling if rather surprising *argument* having among its premises the totality of one's beliefs (I'm here obviously disagreeing with the views that some other theorists have taken on "*bootstrapping arguments*"; see [17] for a seminal presentation of bootstrapping and [30], pp. 48–49, fn 29 for the relevant bit of my own take on these arguments). More importantly, the temptation should definitively vanish in the presence of many *other* examples of the preface paradox in which there is no question that it is the beliefs about particular cases that are among the grounds on which the general belief depends.

held,[13] there would be no room for distinguishing, in the way required for the natural resolution to work, between the belief that every belief one has is true and the belief that any belief one has is true (let alone between the way the world is iff every belief one has is true and the way the world is iff any belief one has is true).[14] Thus, on this resolution, (A$^{\Rightarrow}$E) (and so ($^{\Rightarrow}$C$_{A}$E)) fails.[15,16]

[13]I'll henceforth assume that the relevant principles, suitably understood, cover *restricted* quantification no less than *unrestricted* quantification (see [38] for more details on this issue).

[14]More precisely, there would be no such room *unless* we're envisaging an **LS**ish approach to the natural resolution proposed in the text. For, according to a conception of the *normativity of logical consequence* congenial to **LS**, sometimes one can know that '*P*' entails '*Q*' and nevertheless justifiedly [believe that *P* without believing that *Q*] (see [34] for a development of such conception). However, even on such approach, ($^{\Rightarrow}$C$_{A}$E) would *still* fail, and so there would *still* be the desired difference between anythingness and everythingness. Moreover, the preface paradox also arguably establishes the *stronger* point to the effect that one can justifiedly [believe of every belief one has that it is true *while believing that it is not the case* that every belief one has is true], even if one knows which beliefs one has. And, even according to the just mentioned conception of the normativity of logical consequence congenial to **LS**, if one knows that '*P*' entails '*Q*', one cannot justifiedly [believe that *P* while believing that it is not the case that *Q*]. I'll sketch what an **LS**ish approach would look like in Footnote 20, and I'll note there that, on what might well be the most promising non-transitive approach in the vicinity, (A$^{\Rightarrow}$E) (over and above ($^{\Rightarrow}$C$_{A}$E)) does fail. Similar comments hold in almost their totality for the next two arguments in the text.

[15]Yes, assuming that one can justifiedly believe of every ticket of a fair etc. lottery that it will lose, I'd propose the same view on the *lottery paradox* (see [10], p. 197). A notable feature of the views I'm proposing in this section which would emerge in an especially stark form in the case of the lottery paradox (and which I'm grateful to Sven Rosenkranz for insisting on) is that, in believing [that any ticket will lose while some ticket will win], one believes, roughly, that, no matter which ticket one singles out for consideration, the winning ticket is going to be different from the ticket one has thus singled out. Isn't one thereby *arbitrarily insulating oneself against the counterexample* that one does know to exist? Exactly so!

[16]Speaking about the lottery paradox, after completing a first draft of this paper I became aware of [14], to whose challenging ideas the approach of this paper is broadly congenial. Among other things, Paoli draws on the lottery paradox to argue for a difference between anythingness and everythingness in a way similar to how I've done it in Footnote 15 (although he officially prefers to put the point in terms of an ambiguity in 'everything') and recommends treating the paradox by taking something along the lines of the **LW**ish approach that I'll describe in Footnote 20. While I'm sympathetic to this much, Paoli also seems to require that, in the lottery paradox, *one have no inconsistent justified beliefs*. But, since he is ready to grant that, for every ticket, one justifiedly believes that it will lose, and since he accepts that the contradictory of the justifiedly believed 'Some ticket will not lose' behaves like a big multiplicative conjunction, it would seem that *on Paoli's own view* one does have inconsistent justified beliefs after all. I'm less sanguine about Paoli's two other arguments in favour of the ambiguity of quantifiers. One argument claims that McGee's best-known purported counterexample to *modus ponens* ([13], p. 462) relies on reading 'A Republican will win the election' as *multiplicative* when it occurs as the main antecedent in the conditional premise and as *additive* when it occurs unembedded as the other premise. But it's hard to see that that can adequately deal with the problem, since, even setting aside McGee's other purported counterexamples (in which the target sentence is not quantificational in the first place, but which are admittedly less compelling), the purported counterexample is just as good replacing 'A Republican will win the election' with 'Carter will lose the elections'. The other argument is related, and, playing on Adams' Oswald pairs ([2], p. 70), revolves around the

Secondly (and relatively unsurprisingly given the tight *correspondences* between the *logical behaviour* of *what one justifiedly believes* and that of *what would be the case if something were the case*), consider the less familiar point from discussion of what may be called 'the *party paradox*' (see [29]) to the effect that, for example, it can be the case that, for every friend one has, [if one invited every friend one has to the party, she would come] without its being the case that, if one invited every friend one has to the party, every friend one has would come (a point which I'll henceforth assume without further argument to be correct). The point is to some extent familiar and yet still widely controversial. I think that at least one source of the controversiality consists in the fact that it admittedly seems to border on the *inconsequent* for the invitation to be such that it would cause friend #0 to come, such that it would cause friend #1 to come, such that it would cause friend #2 to come ... without being such that it would cause any *general* fact about the party attendance of one's friends—if the invitation would cause all those things about *particular* friends but would cause no more, it would *mysteriously fail to cause a pattern* that is in fact pinned down by what it would cause (and that however transcends the party attendance of any particular friend). But what could that general fact be if not the fact that every friend one has comes to the party? Here is another way to put the worry. Presumably, when, for every friend one has, the invitation would *consequently* cause her to come to the party, the invitation would thereby cause an *individual, distinctive effect* that goes beyond the effect constituted by the fact that friend #0 comes, beyond the effect constituted by the fact that friend #1 comes, beyond the effect constituted by the fact that friend #2 comes...—the relevant portion of the invitation's counterfactual efficiency does *not reduce to an atomistic plurality of effects each only involving a particular friend*. Moreover, the effect that would be caused by the invitation is *factual*: it is a way in which the world

---

observation that, while Jim's assertion of 'Someone killed Kennedy' based on seeing Kennedy's fatally wounded corpse *supports the inference* to 'If Oswald didn't kill Kennedy, someone else did', Jon's assertion of 'Someone killed Kennedy' based on seeing Oswald shooting does *not* support the same inference. But, again, it's hard to see that that can adequately deal with the problem, since the contrast is just as present replacing 'Someone killed Kennedy' with 'Oswald killed Kennedy' and supposing that Jim recognises the fatal wounds on Kennedy's corpse to have been caused by a rare kind of bullet Oswald is well known for using. Notice also that the last two arguments, but not the argument from the lottery paradox, crucially rely on the logic being *non-monotonic*, at least to the extent that '*a* is *F*' does not entail 'Something is *F*' for one of the senses of 'something'. In our context, that implies in particular that, at least given standard assumptions (shared by Paoli) on the relation between particular and universal quantification (see Footnote 2), those arguments crucially rely on *instantiation principles along the lines of* ($E^{\Rightarrow}C$) *failing* for one of the senses of 'everything'. I myself would share the view that instantiation principles along the lines of ($E^{\Rightarrow}C$) *are not strong enough* to capture the full strength of everythingness (see Footnote 23), but the suggestion that they *fail* strikes me as being at odds with any recognisable notion of everythingness (or, for that matter, anythingness). (Points similar to those made from 'The other argument' onwards apply to the basic argument in favour of the ambiguity of 'or' run for example by [15]; see also [30], pp. 48–49, fn 29.) Thanks to Nissim Francez and Francesco Paoli for help with this fn.

could be. But what, in the example, could that way be if not the way the world is iff every friend one has comes to the party?

This worry admits of a natural resolution that manages to be faithful to its presuppositions while preserving the point forcefully brought out by the party paradox: the general fact that would be caused by the invitation is not the fact that *every* friend one has comes to the party, but the fact that *any* friend one has comes to the party, and the way the invitation's effect is identical with is not the way the world is iff *every* friend one has comes to the party, but the way the world is iff *any* friend one has comes to the party. However, if (A$^{\Rightarrow}$E) held, there would be no room for distinguishing, in the way required for the natural resolution to work, between the fact that every friend one has comes to the party and the fact that any friend one has comes to the party (or between the way the world is iff every friend one has comes to the party and the way the world is iff any friend one has comes to the party). Thus, on this resolution, (A$^{\Rightarrow}$E) (and so ($^{\Rightarrow}$C$_A$E)) fails (see [29] for a unified, **LS**-based treatment of the preface, party and related paradoxes, focussing on different issues from those focussed on in this paper but compatible with the views proposed here).[17]

Thirdly, consider the familiar point from discussion of *free-choice permission* (see [18], p. 7, fn 1, who however also initiates the tradition of what I'll argue in the third point of the next paragraph is too strong a description of free-choice effects) to the effect that, for example, it can be the case that, with the menu of the day, one can have *any* main course, without its being the case that, with the menu of the day, one can have *every* main course. A natural, innovative strategy to account for this contrast is to postulate that the *proposition* that one has any main course is *not at least as strong* as (and is in fact *weaker* than) the *proposition* that one has every main course, and so that *permission* of the former does not entail *permission* of the latter (and is indeed consistent with *prohibition* of the latter). Thus, on this strategy, (A$^{\Rightarrow}$E) (and so ($^{\Rightarrow}$C$_A$E)) fails.

There is however a competing, standard, conservative stategy to account for the contrast, which insists that 'any', just as well as 'every', expresses universal quantification, but then postulates that 'any', contrary to 'every', forces universal quantification to take *wide scope* over the relevant modal element (in our case expressed by 'can').[18] However, this latter strategy is problematic on several counts. First, from a *syntactic* point of view, the strategy is problematic in that it is forced to be restricted to constructions that admit the required scoping out of the universal

[17]The preface, party and related paradoxes all have less-known *single*-premise variations (for which see again [29]). However, it would not seem that worries comparable to those I've discussed arise in such variations (see Footnote 20 for more on single-premise closure principles in our context).

[18]Of course, there are ever so many other—much more sophisticated—competing strategies to account for the contrast and, more generally, for free-choice effects. A critical treatment of these lies however beyond the scope of this paper (which has bigger fish to fry); the focus in the text is justified by the fact that the wide-scope strategy may be particularly salient in our discussion of the difference between anythingness and everythingness (indeed, some audience members seemed to be shocked when they learnt that I reject the wide-scope strategy!).

quantifier, whereas free-choice effects can occur *even if 'any' is in a syntactic island*, as witnessed by 'You can have what anyone has'. Second, from a *semantic* point of view, the strategy is problematic in that it is forced to be restricted to objects that exist independently of the action permitted, whereas free-choice effects can occur *even if one is quantifying over objects that would plausibly only exist if the permitted action were performed*, as witnessed by 'You can have lamb chops and then give any of the leftover bones to the dog'.[19] Third, the strategy is problematic in that it still has the consequence that, for example, 'You can have lamb chops' and 'You can have prawn rice' *both hold even if one cannot have both lamb chops and prawn rice*. Such consequence is already very suspicious in that it reveals that the permission issued by 'You can have lamb chops' must be so extraordinarily weak as to be cancelled by the sheer fact that one has prawn rice. Does one usually issue—or would it usually be sensible to issue—any two permissions such that taking advantage of one cancels the other? Do restaurants systematically violate the **S5**-axiom for deontic modality (and, given the Andersonian-Kangerian reduction (see e.g. [3]), the **S5**-axiom for metaphysical modality)? And the consequence becomes even more problematic by reflecting that 'You can have lamb chops and you can have prawn rice' contrasts with the much more appropriate 'You can have lamb chops or you can have prawn rice', that 'You can have lamb chops and you can have something else' contrasts with the much more appropriate 'You can have lamb chops or you can have something else' and that 'You can have lamb chops and you can have prawn rice' should be expandable into the clearly unacceptable 'You can have lamb chops and also you can have prawn rice'. Indeed, in other examples it is even clearer from a direct inspection that not both permissions hold. For example, if we're sharing a course and two chops c_0 and c_1 are left, I can issue the permission 'You can have any of the two chops' without implying 'You can have c_0 and you can have c_1'. Finally, if a voucher only entitles one to either lamb chops or prawn rice, but not to both, it would be felt as a cheat and indeed as straightforwardly false to write on the voucher both 'Entitles you to lamb chops' and 'Entitles you to prawn rice'. But such statements would both be correct if both 'You can have lamb chops' and 'You can have prawn rice' held. Fourth, as already apparent from the last point, the strategy is problematic in that *analogous free-choice effects occur with 'or'*, as witnessed by 'You can have lamb chops or prawn rice'. The phenomenon of free choice requires *general* insights into *logic* and *semantics* applicable to a *wide* range of natural-language expressions rather than a *specific* postulation about *syntax* concerning *only* 'any' (compare the discussion of related issues in [38]).[20]

---

[19]On the wide-scope reading, the sentence in the text might still come out *vacuously true*, but then the criminal 'You can have lamb chops and then beat the dog with any of the leftover bones' would equally do so. The point in the text is that the wide-scope reading is not a good representation of the *content* of the relevant sentences. Thanks to John Horden for raising this issue.

[20]It will probably not be astonishing to learn that I think that both **LW** and **LS** can offer valuable insights for a deeper understanding of the preface and party paradoxes as well as free-choice permission. Things are more straightforward on an **LW**ish approach. A notable feature of **LW** that is particularly relevant in this respect is that, in it, 'Anything is *F*' (with 'anything' understood to be governed by ($^{\Rightarrow}$C$_{A}$A)) does not entail 'Everything is *F*' (with 'everything' understood as

The last two arguments to the effect that there is a difference between anythingness and everythingness draw directly on considerations from the semantic and vagueness paradoxes and from the theories solving them that I've advocated in

---

per Sect. 21.1). (A bit more in detail, in **LW** $\mathcal{A}$ can be taken to be the usual additive universal quantifier of affine logic.) To get a feel for the workings behind this, consider that, in **LW**, one can use 'Anything is *F*' to infer that, say, object #0 is *F*, but then, because of failure of contraction, one cannot *also* use 'Anything is *F*' to infer that, say, object #1 is *F*. Because, in **LW**, the argument from 'Object #0 is *F*', 'Object #1 is *F*', 'Object #2 is *F*'... to 'Everything is *F*' remains valid, on an **LW**ish approach the views I've proposed on the preface and party paradoxes as well as free-choice permission do imply failures of *multi*-premise closure for justification and counterfactual implication, but not for permission (since, as I've argued, in cases of free-choice permission it is not the case that each of the options is permitted); however, because, in **LW**, the argument from 'Anything is *F*' to 'Everything is *F*' is no longer valid, on an **LW**ish approach the views I've proposed on the preface and party paradoxes as well as free-choice permission do not imply failures of *single*-premise closure for justification, counterfactual implication and permission (see [38] for further details on this approach to free-choice permission and its comparison with another non-contractive approach). Things are less straightforward on an **LS**ish approach. A notable feature of **LS** that is particularly relevant in this respect is that, in it, although 'Anything is *F*' (with 'anything' understood to be governed by ($\Rightarrow$C$_A$A)) entails 'Everything is *F*' (with 'everything' understood as *per* Sect. 21.2), it is still possible, roughly, both to accept 'Anything is *F*' and not to accept 'Everything is *F*'. (A bit more in detail, in **LS** $\mathcal{A}$ can be taken to be constrained by the condition that, under an assignment, the value of $\mathcal{A}\xi\varphi$ is designated (tolerated) iff, under every assignment that is a ξ-variant of the original assignment, the value of φ is designated (tolerated).) To get a feel for the workings behind this, consider that, in **LS**, one can accept 'Anything is *F*', and so, say, 'Object #0 is *F*' or 'Object #1 is *F*', as conclusions of a certain argument with very good premises, in which case those sentences are good enough to serve as *terminal points of acceptance* but, because of failure of transitivity, may not be good enough to serve as *initial points for inferring* that [object #0 is *F* and object #1 is *F*]. Because, in **LS**, the argument from 'Object #0 is *F*', 'Object #1 is *F*', 'Object #2 is *F*'... to 'Everything is *F*' remains valid, on an **LS**ish approach the views I've proposed on the preface and party paradoxes as well as free-choice permission do imply failures of *multi*-premise closure for justification and counterfactual implication, but not for permission (since, as I've argued, in cases of free-choice permission it is not the case that each of the options is permitted); moreover, because, in **LS**, the argument from 'Anything is *F*' to 'Everything is *F*' also remains valid, on an **LS**ish approach the views I've proposed on the preface and party paradoxes as well as free-choice permission do imply failures of *single*-premise closure for justification, counterfactual implication and permission. In this sense, on an **LS**ish approach the *gap* between *justification* and *counterfactual implication* on the one hand and *logic* on the other hand is *widened*, and a new gap *opens up* even between *permission* on the one hand and *logic* on the other hand (even if I should note that, in the framework of **LS**, failures of multi-premise and single-premise closure principles are independently to be expected). In my opinion, an important disadvantage of an **LS**ish approach is that its extension to strengthened versions of the preface and party paradoxes as well as free-choice permission (in which, roughly, failure of accepting 'Everything is *F*' is replaced by acceptance of 'Something is not *F*') faces substantial complications (see Footnote 14). (On the contrary, an **LW**ish approach carries over smoothly to such strengthened versions.) I should add though that the larger family of non-transitive logics to which **LS** belongs (see [23, 37]) contains also *weaker* logics in which the offending arguments (not only the one from 'Anything is *F*' to 'Everything is *F*', but also the one from 'Object #0 is *F*', 'Object #1 is *F*', 'Object #2 is *F*'... to 'Everything is *F*') are no longer valid, and so in which, for better or worse, the gap between *justification*, *counterfactual implication* and *permission* on the one hand and *logic* on the other hand is *bridged* (see [29] for more discussion of this last issue).

Sects. 21.1 and 21.2 respectively.[21] Fourthly, consider the two sentences 'Every L_0 sentence is not true' (λ_0) and 'Every L_1 sentence is not true' (λ_1), where 'L_0' and 'L_1' are two distinct predicates as a matter of fact both applying exactly to λ_0 and λ_1. Suppose that λ_0 is true. Moreover, suppose that either x is λ_0 or x is λ_1. If the former, x is true (since, by supposition, λ_0 is true); if the latter, then again x is true (since, clearly, λ_0 is true iff λ_1 is, and, by supposition, λ_0 is true). Given that both 'x is λ_0' and 'x is λ_1' can be assumed to be effectively classical, 'x is true' can then be taken effectively to follow from 'λ_0 is true' and 'Either x is λ_0 or x is λ_1', and so, by the facts of the matter, 'x is true' can then be taken effectively to follow from 'λ_0 is true' and 'x is an L_0 sentence'. By ($^{\Rightarrow}$C$_{\text{A}}$A), 'Any L_0 sentence is true' follows from 'λ_0 is true'. By (A$^{\Rightarrow}$E) and transitivity of logical consequence, 'Every L_0 sentence is true' follows from 'λ_0 is true'. However, if every L_0 sentence is true, then, by the facts of the matter and a suitable instantiation principle (see Footnote 23), both λ_0 is true and λ_1 is true. And since, by transparency, the facts of the matter and (E$^{\Rightarrow}$C), λ_1 being true entails that λ_0 is not true, the second conjunct of 'λ_0 is true and λ_1 is true' contradicts the first, and so the conjunction is inconsistent. But the conjunction follows from 'λ_0 is true', which, by transitivity of logical consequence, is thus inconsistent. Therefore, by a special version of *reductio ad absurdum*, λ_0 is not true. An analogous reasoning would also establish that λ_1 is not true. But, if λ_1 is not true, then, by transparency, the facts of the matter and contraposition on a suitable generalisation principle (see Sect. 21.5), it is not the case [that λ_0 is not true and λ_1 is not true], and, since λ_0 is not true, by *modus ponendo tollens* it is not the case that λ_1 is not true. Contradiction with the previous conclusion that λ_1 is not true. Since every other step of this argument is valid in **LW**, (A$^{\Rightarrow}$E) (and so ($^{\Rightarrow}$C$_{\text{A}}$E)) has got to fail in **LW**: under the supposition that λ_0 is true, any L_0 sentence is true, but not every L_0 sentence is true.

Fifthly, consider the two sentences β_1 and $\neg\beta_2$, and let '$B_{1,\neg 2}$' be a predicate as a matter of fact applying exactly to β_1 and $\neg\beta_2$. Suppose that $\mathcal{N}$ holds. Then, by *modus ponens*, β_1 holds. Suppose that $\neg\beta_2$ also holds. Moreover, suppose that either x is β_1 or x is $\neg\beta_2$. If the former, x holds (since β_1 follows by *modus ponens* from $\mathcal{N}$); if the latter, then again x holds (since, by supposition, $\neg\beta_2$ holds). Given that both 'x is β_1' and 'x is $\neg\beta_2$' can be assumed to be effectively classical, 'x holds' can then be taken effectively to follow from $\mathcal{N}$, $\neg\beta_2$ and 'Either x is β_1 or x is $\neg\beta_2$',

[21] That is at least my favoured way of seeing things. Another way of seeing things would have it that there is actually a common *big problem* with the theories that I've advocated with regard to the semantic and vagueness paradoxes: namely, that they both require that there be a difference between anythingness and everythingness! Indeed, a related way of seeing things would have it that there is actually a common *big problem* with those theories and with the theories that I've advocated with regard to the preface and party paradoxes as well as free-choice permission: namely, that they all require that there be a difference between anythingness and everythingness! In effect, it lies beyond the scope of this paper to offer detailed arguments in favour of any of these theories. The approach of the paper is rather, by and large, to *assume* such theories and *explore* some of the avenues of inquiry they open up concerning the relation between anythingness and everythingness. Thanks to Sven Rosenkranz for helping me to get clear about this.

and so, by the facts of the matter, ‘x holds’ can then be taken effectively to follow from $\mathcal{N}$, $\neg\beta_2$ and ‘x is a $B_{1,\neg 2}$ sentence’. By ($^{\Rightarrow}$C$_{\text{A}}$A), ‘Any $B_{1,\neg 2}$ sentence holds’ follows from $\mathcal{N}$ and $\neg\beta_2$. By ($^{\Rightarrow}$AE), ‘Every $B_{1,\neg 2}$ sentence holds’ follows from $\mathcal{N}$ and $\neg\beta_2$. However, if ‘Every $B_{1,\neg 2}$ sentence holds’ follows from $\mathcal{N}$ and $\neg\beta_2$, then, by the facts of the matter and a suitable instantiation principle (see Footnote 23), ‘Both β_1 holds and $\neg\beta_2$ holds’ also follows from $\mathcal{N}$ and $\neg\beta_2$. But $\mathcal{N}$ includes ‘It is not the case that both β_1 holds and $\neg\beta_2$ holds’, and so $\mathcal{N}$ is inconsistent with $\neg\beta_2$. Therefore, by a special version of *reductio ad absurdum*, $\mathcal{N}$ entails β_2. This result would be disastrous, as it would in effect amount to showing that, if $\mathcal{N}$ entails β_i (which has here been instantiated with β_1), it also entails β_{i+1} (which has here been instantiated with β_2). Since every other step of this argument is valid in **LS**, ($^{\Rightarrow}$AE) (and so ($^{\Rightarrow}$C$_{\text{A}}$E)) has got to fail in **LS**: under the supposition that $\mathcal{N}$ and $\neg\beta_2$ hold, any $B_{1,\neg 2}$ sentence holds, but not every $B_{1,\neg 2}$ sentence holds.[22]

21.5 ω

If ($^{\Rightarrow}$C$_{\text{A}}$E) fails, what should take its place as a more adequate generalisation principle? An attractive candidate is the *from-the-totality-of-things-to-everything* rule:

(T$^{\Rightarrow}$E) If $\tau_0, \tau_1, \tau_2 \ldots$ are the totality of contextually relevant singular terms of the language, $\varphi_{\tau_0/\xi}, \varphi_{\tau_1/\xi}, \varphi_{\tau_2/\xi} \ldots \vdash \forall\xi\varphi$ holds.

[22]Notice that all the five arguments I've given to the effect that there is a difference between anythingness and everythingness establish, if good, a difference that is strong enough for ($^{\Rightarrow}$C$_{\text{A}}$E) to fail, and so strong enough for envisaging suppositions under which ‘Anything is F’ is true but ‘Everything is F’ is not. In this sense, all the five arguments I've given establish, if good, a *truth-conditional* difference between ‘Anything is F’ and ‘Everything is F’. These arguments are thus in a completely different ballpark, for example, from the arguments given by [7], pp. 81–90 to the effect that there is a difference in the *canonical justificational procedures* associated with ‘*Each* thing is F’ and ‘*Every* thing is F’ respectively: as Fiengo makes clear, the difference he's envisaging, if it exists, only results at best in a difference in *content* (and in *truth conditions for embeddings under verbs of propositional attitudes*). As far as I can tell, the arguments advanced in this paper are neutral with regard to the data and theses offered by Fiengo (some of which do not persuade me, although I must leave those issues for another occasion). Admittedly, the truth-conditional difference I'm envisaging between ‘Anything is F’ and ‘Everything is F’ is *not crudely extensional* in that it does not reflect itself in a difference between which *lists* containing exactly one of each of ‘Object #0 {is F, is not F}’, ‘Object #1 {is F, is not F}’, ‘Object #2 {is F, is not F}’... are compatible with ‘Anything is F’ and which such *lists* are compatible with ‘Everything is F’: both sentences are only compatible with the list ‘Object #0 is F’, ‘Object #1 is F’, ‘Object #2 is F’... But that difference does reflect itself in *different compatibility profiles*: to take a crucial example (see Footnotes 14, 15 and 20), ‘Something is not F’ is compatible with ‘Anything is F’, but it is not compatible with ‘Everything is F’. Thanks to Joe Moore for pushing me on some of these issues.

(T$^{\Rightarrow}$E) has several virtues. For starters, (T$^{\Rightarrow}$E) is *intuitively compelling*, for it is intuitively compelling that the holding of the totality of the contextually relevant instances is *logically sufficient* for the holding of a universal quantification.

Moreover, (T$^{\Rightarrow}$E) sits comfortably with (E$^{\Rightarrow}$C) *no less* than ($^{\Rightarrow}$C$_A$E) does: just as, if anything is F, it certainly follows that a is F (where 'a' is any contextually relevant singular term), so, if the totality of a_0, a_1, a_2 ... are F (where, roughly, 'a_0, a_1, a_2 ...' and its like denote a canonical enumeration of the totality of contextually relevant linguistic entities of a certain kind), it certainly follows that a is F (where 'a' is any contextually relevant singular term). Indeed, (T$^{\Rightarrow}$E) sits comfortably with (E$^{\Rightarrow}$C) *more* than ($^{\Rightarrow}$C$_A$E) does, as, in addition to avoiding all the *bureaucracy* ensuing from the use of eigenvariables, (T$^{\Rightarrow}$E) is (E$^{\Rightarrow}$C)'s *direct converse*, and thus most directly combines with it to *equate logically* the holding of a universal quantification with the holding of the totality of its contextually relevant instances.[23]

Furthermore, (T$^{\Rightarrow}$E) avoids all the problems for ($^{\Rightarrow}$C$_A$E) discussed in Sect. 21.4. Firstly, in the case of the preface paradox, although, for every i, one justifiedly believes that belief #i is true, and although, by (T$^{\Rightarrow}$E), the totality of such beliefs entail that every belief one has is true, it is by no means surprising that one can justifiedly not believe a conclusion that one only knows to follow from too many premises that one justifiedly believes individually: such are the effects of the failure of *multi*-premise closure for justification. Secondly, in the case of the party paradox, although, for every i, the invitation would cause friend #i to come to the party, and although, by (T$^{\Rightarrow}$E), the totality of such facts entail that every friend one has comes to the party, it is by no means surprising that the invitation would not cause a conclusion that only follows from too many premises that the invitation would cause individually: such are the failures of *multi*-premise closure for counterfactual implication (also known as *agglomeration*). Thirdly, in the case of free-choice permission, as I've argued in Sect. 21.4 it is not even the case that each of the

[23]To be a bit more accurate, at least in our target logical environments **LW** and **LS** (T$^{\Rightarrow}$E)'s perfect match is actually a principle guaranteeing that a universal quantification entails *the totality* of its contextually relevant instances *together*, and so something along the lines of the *from-everything-to-the-totality-of-things* left-metarule:

(ET$^{\Rightarrow}$) If $\tau_0, \tau_1, \tau_2 \ldots$ are the totality of contextually relevant singular terms of the language, if $\Gamma, \varphi_{\tau_0/\xi}, \varphi_{\tau_1/\xi}, \varphi_{\tau_2/\xi} \ldots \vdash \Delta$ holds, $\Gamma, \forall\xi\varphi \vdash \Delta$ holds,

whereas (E$^{\Rightarrow}$C), or for that matter the *from-everything-to-a-certain-thing* left-metarule:

(EC$^{\Rightarrow}$) If τ is a contextually relevant singular term of the language, if $\Gamma, \varphi_{\tau/\xi} \vdash \Delta$ holds, $\Gamma, \forall\xi\varphi \vdash \Delta$ holds,

only guarantees that a universal quantification entails *any* of its contextually relevant instances *individually*. (Presumably, a better deductive system than the standard ones that are now available would allow us to formulate the desired instantiation principle—for which *nomen nudum tenemus*: '(E$^{\Rightarrow}$T)'—without smuggling in any form of transitivity of logical consequence, contrary to what (ET$^{\Rightarrow}$) does.) However, this otherwise important distinction among instantiation principles is not relevant for the focus of this paper, and that is why, throughout, I employ the simpler (E$^{\Rightarrow}$C). Thanks to Dave Ripley for comments that led to this fn.

options is permitted, and so it is *not* the case that the totality of the premises of the relevant instance of ($T^{\Rightarrow}E$) are available *in the first place* (thus, contrary to modal lore, *multi*-premise closure for permission does not fail in such case). Fourthly, ($T^{\Rightarrow}E$), contrary to ($^{\Rightarrow}C_{A}E$), *is* valid in the *provably consistent* **LW**. Fifthly, ($T^{\Rightarrow}E$), contrary to ($^{\Rightarrow}C_{A}E$), *is* valid in **LS** in which $\mathcal{N}$ is *provably consistent*.

One routine objection to ($T^{\Rightarrow}E$) as a generalisation principle is that it is *unsound* if *not every object in the contextually relevant domain of discourse is denoted by some contextually relevant singular term of the language*. The objection clearly relies only on the fact that a certain principle becomes unsound once there exists some *lack of correspondence between the contextually relevant singular terms of the language and the objects in the contextually relevant domain of discourse*. Thus, it is only good to the extent that, *conversely*, one can also succesfully object to ($E^{\Rightarrow}C$) as an instantiation principle on the grounds that it is unsound if *not every contextually relevant singular term of the language denotes some object in the contextually relevant domain of discourse*. But the latter is no good objection: ($E^{\Rightarrow}C$) can legitimately be understood to govern only those cases in which every contextually relevant singular term of the language denotes some object in the contextually relevant domain of discourse (cases in which, as it were, the contextually relevant singular terms of the language are "*sound*" with respect to the objects in the contextually relevant domain of discourse). *Just so*, ($T^{\Rightarrow}E$) can legitimately be understood to govern only those cases in which every object in the contextually relevant domain of discourse is denoted by some contextually relevant singular term of the language (cases in which, as it were, the contextually relevant singular terms of the language are "*complete*" with respect to the objects in the contextually relevant domain of discourse).

True, it is also legitimate to ask which generalisation principle should govern those cases—deviating from the logician's simplifying idealisations but approximating the inquirer's complex reality—in which not every object in the contextually relevant domain of discourse is guaranteed to be denoted by some contextually relevant singular term of the language. But, again, it is also *equally* legitimate to ask which instantiation principle should govern those cases—deviating from the logician's simplifying idealisations but approximating the inquirer's complex reality—in which not every contextually relevant singular term of the language is guaranteed to denote some object in the contextually relevant domain of discourse. There is a well-established and, for our purposes, adequate answer to the latter query. The answer has been developed in the field of *free logics* (see [11]), and consists in adding to the logical expressions of the language a *singular objectual predicate* O (such that $O\tau$ is true iff τ does denote some object in the contextually relevant domain of discourse) and then saying that the appropriate instantiation principle for those cases in which not every contextually relevant singular term of the language is guaranteed to denote some object in the contextually relevant domain of discourse is:

($E^{\Rightarrow}C^{O}$) If τ is a contextually relevant singular term of the language, $\forall\xi\varphi, O\tau \vdash \varphi_{\tau/\xi}$ holds.

Just so, one should add to the logical expressions of the language a *plural total predicate* T (such that $T\tau_0; \tau_1; \tau_2 \ldots$[24] is true iff $\tau_0; \tau_1; \tau_2 \ldots$ does plurally denote the totality of objects in the contextually relevant domain of discourse) and then say that the appropriate generalisation principle for those cases in which not every object in the contextually relevant domain of discourse is guaranteed to be denoted by some contextually relevant singular term of the language is:

($T^{\Rightarrow}E^{T}$) If $\tau_0, \tau_1, \tau_2 \ldots$ are the totality of contextually relevant singular terms of the language, $\varphi_{\tau_0/\xi}, \varphi_{\tau_1/\xi}, \varphi_{\tau_2/\xi} \ldots, T\tau_0; \tau_1; \tau_2 \ldots \vdash \forall\xi\varphi$ holds.

Another routine objection to ($T^{\Rightarrow}E$) as a generalisation principle is that it is *epistemically flawed*, in the sense that *someone who accepts its premises may still rationally doubt its conclusion*, since she can rationally doubt that the totality of objects in the contextually relevant domain of discourse have been covered by the premises. This objection fails partly for the reasons explained in the last paragraph. Either ($T^{\Rightarrow}E$) is understood to apply to a case in which every object in the contextually relevant domain of discourse is guaranteed to be denoted by some contextually relevant singular term of the language or it is not. If the latter, a ($T^{\Rightarrow}E$)-inspired approach dictates that the appropriate generalisation principle is actually ($T^{\Rightarrow}E^{T}$), and, because of the presence of the *additional premise* $T\tau_0; \tau_1; \tau_2 \ldots$, the objection evaporates. If the former, the rational doubt appealed to by the objection cannot consist in the doubt that every *object in the contextually relevant domain of discourse* is denoted by some contextually relevant singular term of the language; it can only consist in the doubt that every *contextually relevant singular term of the language* has been covered by the premises.[25] But *that* doubt is a doubt that obviously one cannot rationally have with respect to ($T^{\Rightarrow}E$) *itself*. And, if we assume that one is *competent with the language* (in a sense which implies that one knows which singular terms are part of the language and which are not) and *aware*

[24]Throughout, and roughly, a series of coordinated occurrences of ';' *syntactically* combines *singular* terms into a *plural* term that *semantically plurally* denotes those objects that are *singularly* denoted by some of the combined singular terms.

[25]A different case is one in which, although every object in the contextually relevant domain of discourse is guaranteed to be denoted by some contextually relevant singular term of the language, not every contextually relevant singular term of the language is guaranteed to denote some object in the contextually relevant domain of discourse. In such a case, it might still be objected that ($T^{\Rightarrow}E$) is *epistemically flawed*, in the different sense that *it may not offer a route to knowing a universal quantification*, since one of the premises may not hold (as it may involve a singular term that does not denote any object in the contextually relevant domain of discourse). Such a case has already been treated with regard to the question of finding a valid instantiation principle in the last paragraph in the text. In order to see how the resources introduced there can also be deployed to solve the problem at hand, it is helpful to consider what is, for our purposes, an equivalent version of ($E^{\Rightarrow}C^{O}$):

($E^{\Rightarrow}C^{O\supset}$) If τ is a contextually relevant singular term of the language, $\forall\xi\varphi \vdash O\tau \supset \varphi_{\tau/\xi}$ holds.

($E^{\Rightarrow}C^{O\supset}$) suggests what the proper notion of an instance of a universal quantification generally is in the case in which not every contextually relevant singular term of the language is guaranteed *of the context* (in a sense which implies that one is aware of which singular terms are contextually relevant and which are not), that is also a doubt that one cannot rationally have with respect to any *instance* of (T$^{\Rightarrow}$E) (by which, here and at some other relevant place, I mean something of the form '$\varphi_{\tau_0/\xi}, \varphi_{\tau_1/\xi}, \varphi_{\tau_2/\xi} \ldots \vdash \forall\xi\varphi$', where $\tau_0, \tau_1, \tau_2 \ldots$ satisfy the proviso of being the totality of contextually relevant singular terms of the language).[26]

Granted, if we assume that one is not competent with the language or aware of the context, that becomes a doubt one can rationally have. But, under the same assumption, one can *equally* rationally doubt that any contextually relevant singular term of the language has been covered by the conclusion of an instance of (E$^{\Rightarrow}$C)—indeed, one can *equally* rationally have the corresponding doubts for the instances of many principles of sentential logic (for example, one can equally rationally have the corresponding doubt for the instances of *addition* $\varphi \vdash \varphi \vee \psi$ concerning which, for some reason or other, one doubts that certain expressions in the "added" disjunct are contextually relevant expressions of the language).

---

to denote some object in the contextually relevant domain of discourse, and so leads to a corresponding generalisation principle:

(T$^{\Rightarrow}$E$^{O\supset}$) If $\tau_0, \tau_1, \tau_2 \ldots$ are the totality of contextually relevant singular terms of the language, $O\tau_0 \supset \varphi_{\tau_0/\xi}, O\tau_1 \supset \varphi_{\tau_1/\xi}, O\tau_2 \supset \varphi_{\tau_2/\xi} \ldots \vdash \forall\xi\varphi$ holds.

(Notice that, if it is *also* the case that not every object in the contextually relevant domain of discourse is guaranteed to be denoted by some contextually relevant singular term of the language, it will not do simply to add as further premise $T\tau_0; \tau_1; \tau_2 \ldots$, since that premise may not hold (as it may involve a singular term that does not denote any object in the contextually relevant domain of discourse). For such cases, short of going for a "monotonic" reinterpretation of T, an appropriate generalisation principle is the result of deleting 'the totality of' in (T$^{\Rightarrow}$E^{T}).) By (E$^{\Rightarrow}$C$^{O\supset}$) (or (E$^{\Rightarrow}$T$^{O\supset}$), see Footnote 23), if a universal quantification holds the totality of the premises of the relevant instance of (T$^{\Rightarrow}$E$^{O\supset}$) hold, which solves the problem at hand. This is perhaps the most appropriate place for noting that the objection from *epistemic* flawedness under discussion in the text and in this fn has a variation in terms of *metaphysical* flawedness (to the effect that *the premises of* (T$^{\Rightarrow}$E) *could hold without the conclusion holding*), and that everything I'm saying applies *mutatis mutandis* to such variation. In particular, notice that, on such variation, the problem at hand becomes something like the vexed problem of *finding truth makers for true universal quantifications* (see [4], pp. 196–201), and that (T$^{\Rightarrow}$E$^{O\supset}$) is then also *a solution to that problem that respects the necessitation constraint on truth making*: what makes true and necessitates 'Everything is F' is the collection of facts that, if object #0 is a thing, it is F, that, if object #1 is a thing, it is F, that, if object #2 is a thing, it is F...

[26]Under these assumptions, which determine what is arguably in many respects the theoretically central case, everythingness is *displayed* in the language and need not be *said* (although it can), just as, for example, in many languages *bothness* (i.e. totality of cardinality 2, cf French *tous les deux* and Italian *tutti e due*) is displayed and need not be said (in the relevant contexts, 'Both are engineers' follows from 'Jim is an engineer' and 'Jon is an engineer' without need of a further premise to the effect Jim and Jon are the totality of the relevant objects). In this regard, the argument in the text, as well as other things in this paper, is broadly congenial to the views on generality of [21].

Yet another routine objection to ($T^{\Rightarrow}E$) is that it is *infinitary*. Most naturally understood, the claim just mentioned is actually incorrect, since, although for our purposes conveniently formulated in an informal way with the poetic trick of the "dots of infinity", in a more formal setting ($T^{\Rightarrow}E$) would naturally be spelt out as a sentence of finite length. But, for the sake of argument, let's focus mainly on the less natural understanding under which the claim just mentioned is correct—namely, the fact that, for some languages and contexts, ($T^{\Rightarrow}E$) has infinitary instances. It is absolutely not clear why that is supposed to be a problem for a *logical* principle *qua logical* principle—why should logical consequence not have infinitary features? The objection would seem to presuppose at least two things: that a *logical* principle must also be an effective principle of *inference*, or, more accurately, of *human* inference, and that an infinitary principle like ($T^{\Rightarrow}E$) cannot be an effective principle of human inference. As for the former presupposition, it is not very clear why logical consequence, an absolutely *objective* subject matter comparable to, say, mathematics, should be sensitive to what is an effective inference by the standards of a certain population on a certain planet in the universe. But, for the sake of argument, let's focus on the latter presupposition that ($T^{\Rightarrow}E$) is not an effective principle of human inference.

That presupposition is arguably incorrect in several respects. To see this, it will be useful to consider first contexts in which the relevant domain of discourse is *finite*. Consider a context in which the relevant domain of discourse is constituted by Alf, Bill, Charlie... and Zach (26 objects), and the question is whether everyone has come to the class. One checks and gets to know that Alf has come to the class, that Bill has come to the class, that Charlie has come to the class... and that Zach has come to the class. From this, one arrives at the conclusion that everyone has come to the class. The by far most natural *rationalisation* of one's reasoning has it that one has inferred from the 26 premises 'Alf has come to the class', 'Bill has come to the class', 'Charlie has come to the class'... and 'Zach has come to the class' the conclusion 'Everyone has come to the class' using the relevant finitary instance of ($T^{\Rightarrow}E$). I suppose that the alternative rationalisation that does not appeal to ($T^{\Rightarrow}E$) is meant to go *via* ($^{\Rightarrow}C_AE$). But what would that rationalisation be? As far as I can tell, at a first pass it would have it that, if *x* is an arbitrary person, one accepts the additional assumption 'Either *x* is Alf or *x* is Bill or *x* is Charlie... or *x* is Zach', that, by indiscernibility of identicals and reasoning by cases, one infers from that and the original 26 premises '*x* has come to the class' and that, by ($^{\Rightarrow}C_AE$), one concludes 'Everyone has come to the class' from all this. But, obviously, that would not be a correct application of ($^{\Rightarrow}C_AE$), since one of the assumptions in question contains '*x*' free. There are various strategies with which one could try to patch this up, which I'll list in what I suppose to be a climax of plausibility.

A first strategy would be to *add, by the deduction theorem, the intermediate conclusion* 'If either *x* is Alf or *x* is Bill or *x* is Charlie... or *x* is Zach, *x* has come to the class', which would allow one to discharge the offending assumption for then concluding, by a now correct application of ($^{\Rightarrow}C_AE$), to 'Everyone is such that, if either she is Alf or she is Bill or she is Charlie... or she is Zach, she has come to the class', which we can in turn assume to be tantamount to the more natural

'Everyone who either is Alf or is Bill or is Charlie...or is Zach has come to the class'. In addition to imposing *yet further complexity*, as it stands this strategy faces the glaring problem of not getting to the *desired* conclusion. One needs to add as additional premise something along the lines of 'Everyone who either is Alf or is Bill or is Charlie...or is Zach has come to the class only if everyone has come to the class', *making in effect explicit* the restriction on the contextually relevant domain of discourse and *imposing even more complexity*. I take it that the resulting baroque ($^{\Rightarrow}C_AE$)-based rationalisation has little to recommend it compared to the noble simplicity and quiet majesty of the ($T^{\Rightarrow}E$)-based rationalisation.

A second strategy would be to *modify the other basic principles of the underlying logic*, singling out a special category of "*domain-of-discourse-specifying sentences*" (presumably those sentences of the form 'Either ξ is τ_0 or ξ is τ_1 or ξ is τ_2...or ξ is τ_i') and letting such sentences be *logical truths*, thus turning the required application of ($^{\Rightarrow}C_AE$) into a correct one. Obviously, one can't really let all domain-of-discourse-specifying sentences be logical truths at the same time; one can't even really do so by just restricting to domain-of-discourse-specifying sentences containing only contextually relevant singular terms (for, by transitivity of identity, 'x is Alf' and 'x is Bill' entail 'Alf is Bill'); one can't even really do so by just somehow restricting all this to applications of ($^{\Rightarrow}C_AE$) (for, if 'x is Alf' is a logical truth, by ($^{\Rightarrow}C_AE$) so is 'Everyone is Alf'); in all these cases, the resulting modified logic would be unsound. What one really needs to do is to add the restriction that, in order to count as a logical truth in a context, a domain-of-discourse-specifying sentence must be true in the context. Well, in fact, since a domain-of-discourse-specifying sentence is essentially open the restriction must really be to the effect that such sentence *is satisfied by every object in the domain of discourse relevant for the context*. This strategy does *decrease* a little bit the complexity introduced by the first strategy, but it is nevertheless beset by a host of problems. Firstly, the class example would still seem natural if we replace 'has come to the class' with 'either is Alf or is Bill or is Charlie...or is Zach', but the strategy applied to this modified example would *short-circuit*, since the resulting rationalisation would boil down to simply inferring 'Everyone either is Alf or is Bill or is Charlie...or is Zach' from 'Either x is Alf or x is Bill or x is Charlie...or x is Zach', thereby completely missing the crucial feature of the example consisting in the fact that *one reasons to a universal quantification from its instances* (more accurately, it would boil down to either that or doing something involving the dumb inference of 'Either x is Alf or x is Bill or x is Charlie...or x is Zach' from itself and 'Either Alf is Alf or Alf is Bill or Alf is Charlie...or Alf is Zach', 'Either Bill is Alf or Bill is Bill or Bill is Charlie...or Bill is Zach', 'Either Charlie is Alf or Charlie is Bill or Charlie is Charlie...or Charlie is Zach'...and 'Either Zach is Alf or Zach is Bill or Zach is Charlie...or Zach is Zach'). Secondly, the strategy only saves the generalisation principle in question at the cost of *adding to the other basic principles of the underlying logic*. Thirdly, such addition consists in implausibly *letting open sentences count as basic logical "truths"*, thereby introducing all the *difficulties concerning the interpretation of eigenvariables* into the very foundations of logic (notice that this cannot be remedied by the quantificational understanding of eigenvariables proposed in Sect. 21.3 on pain of making the whole strategy virtually indistinguishable from the third strategy considered in the next paragraph; I'll give a taste of said difficulties during the discussion of a different example introduced in the sixth next paragraph). Fourthly, the strategy has it in effect that 'Everyone has come to the class' *follows from just* 'Alf has come to the class', 'Bill has come to the class', 'Charlie has come to the class'...and 'Zach has come to the class' (since these are the only assumptions on which that conclusion depends); but that would seem to defeat the whole purpose of the exercise, since it makes the use of the relevant instance of ($T^{\Rightarrow}E$) unobjectionable. Fifthly, whenever similarly reasoning under a *different supposition about the contextually relevant domain of discourse*, the strategy would have it that a different open sentence is a basic logical "truth", thereby rebarbatively treating as a logical truth what is just an assumption in all but name.[27,28]

A third strategy would be to *start with the additional premise* 'Everyone either is Alf or is Bill or is Charlie...or is Zach', from which, by ($E^{\Rightarrow}C$) one could infer 'Either *x* is Alf or *x* is Bill or *x* is Charlie...or *x* is Zach', thus avoiding

[27]It might be worried that also the ($T^{\Rightarrow}E$)-based rationalisation is committed to "weird" logical truths, since it licences the derivation of 'Everyone either is Alf or is Bill or is Charlie...or is Zach' from 'Either Alf is Alf or Alf is Bill or Alf is Charlie...or Alf is Zach', 'Either Bill is Alf or Bill is Bill or Bill is Charlie...or Bill is Zach', 'Either Charlie is Alf or Charlie is Bill or Charlie is Charlie...or Charlie is Zach'...and 'Either Zach is Alf or Zach is Bill or Zach is Charlie...or Zach is Zach'. Even setting aside the important distinction between *derived* logical truths (which the ($T^{\Rightarrow}E$)-based rationalisation is at most committed to) and *basic* ones (which the strategy in question is committed to, and on which the second, third and fifth problems focus), the points made in the discussion of the objection from epistemic flawedness suffice to undermine this worry. For either the class example is understood to be a case in which every contextually relevant singular term of the language is guaranteed to denote some object in the contextually relevant domain of discourse (as it emphatically is in the text) or it is not. (I don't think that, naturally understood, the worry has really much to do with a case in which not every object in the contextually relevant domain of discourse is guaranteed to be denoted by some contextually relevant singular term of the language: it is not as though one can only identify Alf, Bill, Charlie...and Zach and one has then the *generic* worry that there could be someone else in the contextually relevant domain of discourse without being able to specify who this could be; rather, one can identify another (existing or non-existing) possible person as well, say, Ann, and one has then the *specific* worry that Ann could be in the contextually relevant domain of discourse. Anyways, what I'd say about the innatural version of the worry will be clear from what I'll say about its natural version.) If the former, it should just be expected that 'Everyone either is Alf or is Bill or is Charlie...or is Zach' counts as a logical truth. If the latter, we've seen that a ($T^{\Rightarrow}E$)-inspired approach dictates that the appropriate generalisation principle is actually ($T^{\Rightarrow}E^{O\supset}$), which will require the *non-logical* premise 'If Ann is a thing, either Ann is Alf or Ann is Bill or Ann is Charlie...or Ann is Zach' (which holds in the context of the class example because the *non-logical* negation of its antecedent—that is, 'Ann is not a thing'—holds in that context) for deriving 'Everyone either is Alf or is Bill or is Charlie...or is Zach', so that this will no longer count as a logical truth.

[28]When thinking about the second to fifth problems, it is useful to keep in mind that and why it is not possible to revise the strategy in question and *modify* ($^{\Rightarrow}C_AE$) *itself* rather than some other basic principles of the underlying logic, again singling out a special category of domain-of-discourse-specifying sentences but this time taking the bull by the horns and straightforwardly allowing that, in ($^{\Rightarrow}C_AE$), *τ may occur free in a domain-of-discourse-specifying assumption*. For,

relying on it as an assumption or as a logical "truth". This strategy *increases* a little bit the complexity introduced by the second strategy, and it is also beset by a host of problems. Firstly, the first problem affecting the second strategy affects *in an even more acute form* the strategy in question. Secondly, the strategy faces the embarrassing question of *why* one accepts 'Everyone either is Alf or is Bill or is Charlie...or is Zach'. It is implausible that one accepts it *non-inferentially*. For one thing, one *needs to think a bit* about that sentence before accepting it, which is usually a good indication that one does not accept the relevant sentence non-inferentially. Moreover, that is such a *structurally complex* and *logically strong* sentence that can hardly be supposed to be non-inferentially accepted. However, the by far most natural *inferential* route starts with the structurally simple and logically weak 'Alf is Alf', 'Bill is Bill', 'Charlie is Charlie'... and 'Zach is Zach' to get, by addition, to the intermediate conclusions 'Either Alf is Alf or Alf is Bill or Alf is Charlie... or Alf is Zach', 'Either Bill is Alf or Bill is Bill or Bill is Charlie... or Bill is Zach', 'Either Charlie is Alf or Charlie is Bill or Charlie is Charlie... or Charlie is Zach'... and 'Either Zach is Alf or Zach is Bill or Zach is Charlie... or Zach is Zach', from which, by the relevant instance of ($T^{\Rightarrow}E$), it concludes to 'Everyone either is Alf or is Bill or is Charlie... or is Zach'. More generally, it's hard to see how an inferential acceptance of 'Everyone either is Alf or is Bill or is Charlie... or is Zach' could be grounded without appeal to some instance of ($T^{\Rightarrow}E$). If this is correct, the strategy in question could not do away with ($T^{\Rightarrow}E$) after all. (Notice that this problem would also seem to affect to some very substantial extent the second strategy.) Thirdly, the strategy can only be applied in those cases in which *the language contains the resources* to define something like disjunction and identity, while one can certainly reason about everythingness in the absence of such notions. Fourthly, I'll argue in the third next paragraph that analogous examples exist that involve a contextually relevant *infinite* domain of discourse, for which friends of ($^{\Rightarrow}C_AE$) will be wary of pursuing an analogous strategy. (Notice that the third and fourth problems also affect the first and second strategy.)

I think it's fair to conclude, to a first approximation, that, in many ordinary contexts in which the relevant domain of discourse is finite, ($T^{\Rightarrow}E$) has a much better claim than ($^{\Rightarrow}C_AE$) to be the generalisation principle that is used in human inference (the approximation concerns the characterisation of what the relevant range of contexts really is; the one just given is good enough for the time being, but I'll try to improve on it in the third next paragraph). This conclusion is already very significant, because, given the strength of the argument in favour of ($T^{\Rightarrow}E$) in the case of contexts in which the relevant domain of discourse is finite, it would be surprising if an *altogether different principle* were used in the case of contexts in which the relevant domain of discourse is infinite: what seems to

setting aside other, comparatively minor issues, the resulting modified logic would not even be sound. Suppose for *reductio ad absurdum* that x is Alf. Then, by the modified version of ($^{\Rightarrow}C_AE$), everyone is Alf. But someone is not Alf. Contradiction. Therefore, by *reductio ad absurdum*, x is not Alf, and so, by ($^{\Rightarrow}C_AE$) (modified or not), everyone is not Alf (depending only on 'Someone is not Alf'). But, although someone is not Alf, not everyone is not Alf.

be *uniform* reasoning about everythingness would implausibly be *split* into two altogether different kinds of reasoning, with altogether different principles being used according to the size of the contextually relevant domain of discourse. And, even if that were so, the point would still remain that we should include ($T^{\Rightarrow}E$) in our logic as the generalisation principle to be used in the case of contexts in which the relevant domain of discourse is finite.[29] Let's consider however in some detail the opposite kind of context in which the relevant domain of discourse is infinite, probably *the* kind of context that is supposed to motivate ($^{\Rightarrow}C_AE$) over ($T^{\Rightarrow}E$) as a principle of human inference.

It is usual to claim that infinitary instances of ($T^{\Rightarrow}E$) *are not usable by humans*. In a natural sense that will become clearer as this section progresses (and that will be enforced in this and the next two paragraphs), that is definitely correct. In the same sense, however, for a primitive $\mathbf{ack}(1,000,000)$ary[30] conjunctive operator $\&_{\mathbf{ack}(1{,}000{,}000)}$, the rule:

$$(\&_{\mathbf{ack}(1{,}000{,}000)}\text{-INTRO})\ \varphi_0, \varphi_1, \varphi_2 \ldots, \varphi_{\mathbf{ack}(1{,}000{,}000)-1} \vdash \&_{\mathbf{ack}(1{,}000{,}000)}(\varphi_0, \varphi_1, \varphi_2 \ldots, \varphi_{\mathbf{ack}(1{,}000{,}000)-1})$$

is not usable by humans either, but that does not in the least detract from ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO)'s claim to be the adequate introduction principle for $\&_{\mathbf{ack}(1{,}000{,}000)}$ *even from the point of view of human inference*. Someone will be tempted to reply that ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO), but not the infinitary instances of ($T^{\Rightarrow}E$), is *in principle* usable by humans. If such reply simply consists in holding that ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO), but not the infinitary instances of ($T^{\Rightarrow}E$), is usable by *finite extensions* of humans, it fails to *explain* what is *wrong* with the infinitary instances of ($T^{\Rightarrow}E$) that is not wrong with ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO). The *objective* facts simply are that "**ack**(1,000)-bounded" (you know what I mean!) extensions of humans cannot use either ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO) or the infinitary instances of ($T^{\Rightarrow}E$), finite extensions of humans can use ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO) but not the infinitary instances of ($T^{\Rightarrow}E$), infinite extensions of humans can use both ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO) and the infinitary instances of ($T^{\Rightarrow}E$)—how are such facts supposed to put ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO) on the right side and the infinitary instances of ($T^{\Rightarrow}E$) on the wrong side?

The reply thus often presupposes that, while we can form a *good conception* of an arbitrary finite extension of a human, we cannot do so for an infinite extension of a human. But it's hard to see that there is any ultimately compelling interesting sense in which the latter is true without the former being false: it may be granted that our conception of an infinite extension of a human is so sketchy and wanting as to fail to meet whatever is the operative standard of goodness, but then also our conception of an "**ack**(1,000,000)fold" (you know what I mean!) extension of a human is pretty much such (a point that has repeatedly been made by *strict finitists*, see [5]; on these topics, I myself have in particular been influenced by [20], pp. 192–221, [22]).

[29]Thanks to Julien Murzi for questions that helped to bring out these points.

[30]Where **ack** is a suitable 1ary version of the Ackermann function (see [1] for the original 3ary version).

If ($\&_{\mathbf{ack}(1,000,000)}$-INTRO) and **ack**(1,000,000)ary instances of ($T^{\Rightarrow}E$) can be vindicated as adequate even from the point of view of human inference by considering how an **ack**(1,000,000)fold extension of a human would reason (for example, in **ack**(1,000,000)fold analogues of the 26fold class example)—and they can—then infinitary instances of ($T^{\Rightarrow}E$) can be vindicated as adequate even from the point of view of human inference by considering how an infinite extension of a human would reason (for example, in an infinitary analogue of the 26fold class example).

Having got so far, it should also become clear that the point brought out by the class example does not really rely on the *finitude* of the contextually relevant domain of discourse (although it is more straightforward to make in that case): also in an infinitary analogue of the 26fold class example, and for essentially the same reasons presented with respect to the 26fold class example, ($T^{\Rightarrow}E$) has a much better claim than ($^{\Rightarrow}C_AE$) to be the generalisation principle that is used in human inference. What the point brought out by the class example does really rely on is the fact that, in the example, *one reasons to a universal quantification from its instances*, and that one does so not by having *a single ground for all instances*, but by having *different grounds for different instances*. Clearly, that can happen quite independently of the size of the contextually relevant domain of discourse.[31] True, while *informally clear enough*, with a modicum of ingenuity the distinction between having a single ground for all instances and having different grounds for different instances *can formally be muddled*, and it is especially easy to do so in those cases in which the contextually relevant domain of discourse is finite, as witnessed by the formal tricks devised by the three strategies we've examined. But, while such tricks may suffice, in those cases, to deliver a ($^{\Rightarrow}C_AE$)-*based argument with true premises* (well, the second strategy actually only delivers a ($^{\Rightarrow}C_AE$)-based argument with premises satisfied by every object in the contextually relevant domain of discourse), they fail to provide a plausible *rationalisation of our reasoning*, or so I've argued (in the other direction, an at least decent ($^{\Rightarrow}C_AE$)-based rationalisation of our reasoning will be available even in those cases in which the contextually relevant domain of discourse is finite, as long as one has a single ground for all instances, since such ground will put one in a position to establish in a natural way the input for an application of ($^{\Rightarrow}C_AE$)).

Beyond the question as to whether we *could possibly use* infinitary instances of ($T^{\Rightarrow}E$), there remains the question as to whether we *do actually use* some such instances. I'd like to argue that we do. Consider for example our acceptance of:

(SUCC) Every natural number has a successor.

I'll assume that either we accept (SUCC) non-inferentially or we accept it either in virtue of an application of ($^{\Rightarrow}C_AE$) or in virtue of an application of ($T^{\Rightarrow}E$).

[31] And, having got so far, it should also become clear that the point brought out by the class example does not really rely on *restrictions* on the contextually relevant domain of discourse (although it is more straightforward to make in that case).

There are good reasons for thinking that we don't accept (SUCC) non-inferentially. Firstly, one *needs to think a bit* about (SUCC) before accepting it, which, again, is usually a good indication that one does not accept the relevant sentence non-inferentially. Secondly, (SUCC) is such a *structurally complex* and *logically strong* sentence that, again, can hardly be supposed to be non-inferentially accepted. Thirdly, if we accepted (SUCC) non-inferentially, it would be mysterious why we could *not* so accept many other sentences of the standard language of arithmetic having the *same logical form*. Fourthly, if we accepted (SUCC) non-inferentially, we'd have a reason depending on it for accepting '0 has a successor'. However, we clearly have another reason for accepting '0 has a successor', namely that we know it to follow from '1 succeeds 0' (which in turn we plausibly have a non-inferential reason to accept). But it does not feel as though we have *two such independent reasons* for accepting '0 has a successor' (for one thing, *defeats* of '1 succeeds 0' seem to defeat '0 has a successor' as well, which is usually a good indication that every reason one has for accepting the weaker sentence depends on reasons one has for accepting the stronger sentence).

There are also good reasons for thinking that we don't accept (SUCC) in virtue of an application of ($^{\Rightarrow}C_AE$).[32] I'll assume that, on the most plausible way of spelling this out, that would require accepting '*x* has a successor' as a logical "truth", as *per* the second strategy discussed in relation to the class example (the third strategy being clearly not applicable). That is however subject to the third problem identified there. That problem has often been addressed by proposing that acceptance of '*x* has a successor' be cashed out in terms of acceptance of '*An arbitrary natural number* has a successor', henceforth assuming an understanding of this different from the quantificational understanding proposed in Sect. 21.3 (which is not going to be available in this particular case, since '*x* has a successor' is supposed to be precisely the point of entry to the relevant general fact). In this kind of context, '*An F* is *G*' is naturally understood as being tantamount to '*Every F* is *G*', but, on this proposal, 'An arbitrary natural number has a successor' cannot be tantamount to '*Every arbitrary natural number* has a successor' (as nothing would then have been done to advance our understanding of how we infer to the relevant general facts). The proposal is thus best construed as being to the effect that acceptance of '*x* has a successor' be cashed out in terms of acceptance of '*The arbitrary natural number k* has a successor', where '*k*' is understood to denote a particular, given object.

I submit that, according to this view, *k* is *different from* 0, 1, 2…, for, if it were identical to any of those, then '*k* has a successor' would either boil down to '0 has a successor', or it would boil down to '1 has a successor', or it would boil down to '2 has a successor'…, and it's hard to see how a one-case (non-mathematical) induction from any of 0, 1, 2... is more justified in this case than in many other cases (for one thing, we know of each of 0, 1, 2…that it is representative of the totality of

[32] Thanks to Julien Murzi, Francesco Paoli, Dave Ripley and Lionel Shapiro for pressing me on this.

natural numbers in some respects but not in many other ones).[33] Yet, I also submit that, according to the view under consideration, *k* is a *natural number*, for, if it were not a natural number, it would be ludicrous to assume (as it needs to be assumed about arbitrary natural numbers if they are to serve their intended role in reasoning) that it has a successor, that it is either odd or even, that it obeys the laws of addition etc. Unfortunately, the resulting combination of claims (i.e. that *k* is different from 0, 1, 2... and that it is a natural number) *contradicts the standard conception of natural numbers*, thereby falsifying the *second-order induction axiom* (considering the property of either being 0 or being 1 or being 2...).

Moreover, it is doubtful that we have any good conception of how to establish '*k* has a successor', for it is doubtful that we have any good conception of what "*the successor of k*" could be in the first place. It cannot be itself an arbitrary natural number *l*, for otherwise it would follow that an arbitrary natural number (*l*) is the successor of an arbitrary natural number (*k*), and so presumably that every natural number is the successor of every natural number.[34] But it cannot be any of 1, 2, 3... either, for otherwise it would follow, by injectivity of succession, that *k* itself is one of 0, 1, 2... (contrary to what I've argued in the last paragraph) and, even worse, it would also follow that some of 1, 2, 3... is the successor of every natural number. So what could it be?

Furthermore, not only does the view under consideration lead both to certainly false mathematical claims and to deeply obscure metaphysical speculations, it also would not seem to serve its epistemological purpose. For, even granting that we have

[33] Some might think that one should conclude instead that *k* is neither determinately identical with *nor determinately different from* any of 0, 1, 2... (For good measure, one might add the stock gloss that it is determinate that either *k* is identical with 0 or *k* is identical with 1 or *k* is identical with 2... The gloss generates a version of ω-inconsistency and does nothing to address the points that I'm going to make—in fact, it might even exacerbate the one about its not being determinate that '*k* has a successor' does not boil down to '0 has a successor'.) Such a view burdens the theory of arbitrary objects with a commitment to indeterminate identity. The commitment is well known for being one of dubious coherence (see [6]; but see [31] for a recent defense of such coherence in the framework of $\mathcal{N}_{\mathbf{LS}}$). Worse, the view does nothing to block any of the relevant arguments in the text: since it is not determinate that *k* is different from, say, 0, it is not determinate that '*k* has a successor' does not boil down to '0 has a successor'; since *k* is neither determinately identical with 0 nor determinately identical with 1 nor determinately identical with 2..., the second-order induction axiom is false (considering the property of either being determinately identical with 0 or being determinately identical with 1 or being determinately identical with 2...).

[34] One might think that the presumed consequence does not follow at least in the theory of arbitrary objects of [8], on which *l* would be considered a "*dependent*" arbitrary object. However, this thought relies on the mistaken assumption that, in Fine's theory, an arbitrary object like *k* has a successor, which, in the relevant (i.e. *literal*) sense, it does not—indeed, in the theory, in the relevant sense, *k* is not a natural number in the first place! Fine does postulate a *non-literal* sense in which '*k* is a natural number' is true, namely the sense in which *it does not talk about k but about natural numbers instead*, and in which it is tantamount to something along the lines of 'Every natural number is a natural number', and it is in that sense that, in the theory, the sentence serves in its intended role in reasoning. But that implies that Fine's theory is unsuitable for substantiating the idea that we accept (SUCC) in virtue of accepting something *about k*.

somehow managed to establish 'k has a successor', it is extremely unclear how big a step that is towards establishing 'Every natural number has a successor'. On the face of it, that is just one more apparently unjustified *one-case induction*. Going back to a theme already surfaced in the second last paragraph, I presume that such induction is supposed to be better than the one-case induction from, say, '0 has a successor' because k, but not 0, is an *arbitrary* natural number. But the intuitive conception of an arbitrary natural number, as the result of *taking a standard natural number and stripping it of its "specific" properties*, is—in its reliance on an unexplained notion of specificity—too vague to support the inference. The inference would indeed be adequately supported if one understood 'specific' as 'not exemplified by every standard natural number'. As a mere stipulation, that however would seem to put k back beyond our immediate epistemic reach, placing it at the same epistemic distance as facts about everythingness. Moreover, a full-blooded stipulation of this kind is inconsistent, since the property of being arbitrary is exemplified by k but it is not exemplified by every (indeed, any) standard natural number.[35] It thus remains extremely unclear how the inference from 'k has a successor' to 'Every natural number has a successor' could be supported. And, even if it could to some extent be supported, the luminous certainty of (SUCC) arguably outruns any support that can be offered by a one-case induction on such a shady case as k.[36]

There are thus good reasons for thinking both that we don't accept (SUCC) non-inferentially and that we don't accept it in virtue of an application of ($^{\Rightarrow}C_AE$). I thus conclude that we do after all accept (SUCC) in virtue of an application of ($T^{\Rightarrow}E$), and so that we do actually use infinitary instances of ($T^{\Rightarrow}E$). But just how can we use such instances? Distinguish two ways in which one can use a

[35]Don't say that one can easily screen off the offending properties as being "*external*" (paradigm example: being thought about by me on 06/12/2013) rather than "*internal*" (paradigm example: being even). It is extremely unclear why being arbitrary should count as external rather than internal. Moreover, being different from 0 (or not being determinately identical with 0) is certainly internal, it is exemplified by k but it is not exemplified by every standard natural number.

[36]An alternative proposal for how we might accept (SUCC) that has often been put forth to me and that, while still being congenial to the spirit of ($^{\Rightarrow}C_AE$), at least has the merit of shunning arbitrary objects would be to say that, for a *particular standard* natural number, say, 35 (it does not really matter which, as it will soon be apparent), one somehow manages to establish that it has a successor, for then, reflecting that *one has established that claim only relying on properties of* 35 *that every natural number exemplifies (plus valid principles)*, inferring (SUCC). Setting aside the obvious question of how one knows that the relevant properties are exemplified by *every* natural number, I submit that this would be a good reason for accepting (SUCC) only if it were the application of a *generally reliable method* for inferring a universal quantification. Unfortunately, the method in question is bankrupt. For example, I know that 35 exemplifies the property of being self-identical, from which it follows, by λ-conversion, that 35 is identical with 35. I've established that 35 is identical with 35 only relying on a property of 35 (the property of being self-identical) that every natural number exemplifies (plus the valid principle of λ-conversion). Thus, it should follow that every natural number is identical with 35. Obviously, this fails as a counterexample to ($^{\Rightarrow}C_AE$) itself, since, upon formalisation of the example, τ would occur free in φ, but there would seem to be nothing in the informal thought behind the proposal that could justify such restriction. Thanks to Giovanni Merlo for discussion of this alternative.

principle. In a first, *non-reflective* way, one actively considers *of each* premise that it holds, and, on the basis of one's understanding of the premises and of the target conclusion, goes from the former to the latter, without necessarily actively considering the validity of the principle itself or of its relevant instance. In a second, *reflective* way, one actively considers the validity of the relevant instance of the principle, and, actively considering *that the* premises hold together, derives the conclusion, without necessarily actively considering of each premise that it holds. To make the contrast vivid, compare the way in which, in the class example, having personally gone through the students present at the class one by one, one may actively consider that Alf has come the class, that Bill has come the class, that Charlie has come the class...and that Zach has come the class, and, on the basis of one's understanding of these premises and of the target conclusion that everyone has come the class, go from the former to the latter without actively considering the validity of (T$^{\Rightarrow}$E) itself or of its relevant instance, with the way in which, having listened to the secretary's listing the students present at the class (attentively enough to know that the totality of students have been mentioned, but casually enough not to have actively considered each student), one may actively consider the validity of the relevant instance of (T$^{\Rightarrow}$E), and, actively considering that the premises hold together, derive the conclusion that everyone has come to the class without actively considering that Alf has come the class, that Bill has come the class, that Charlie has come the class...and that Zach has come the class.

It is in this second, reflective way that I think it's plausible that we do actually use infinitary instances of (T$^{\Rightarrow}$E) in coming to accept, for example, (SUCC). But just in what sense do we actively consider that the premises of such instances hold together? On this, let me conjecturally offer a promissory sketch of a speculative account tentatively revolving around five hypothetical elements. Firstly, let's shift our focus from the *event of active consideration* to the *underlying attitudinal state*, which for the time being we can neutrally express with '*A*' and its like. Notice then that what I mean by, say, '*A*ing that '0 has a successor' holds' is something very informal and pre-theoretic to the effect that one *A*s that 0 has a successor: emphatically, considering that a premise holds is an attitude (at least typically) directed towards *non-linguistic* rather than *linguistic* matters.

Secondly, the typically non-linguistic feature of the contents of *A*ing highlighted in the last paragraph makes *believing* unsuitable for being *A*ing since, henceforth setting aside any sort of semantic ascent in the contents of belief, believing that the premises hold together would seem to *require believing each premise individually*, which one certainly does not in many cases in which it is plausible that we do actually use infinitary instances of (T$^{\Rightarrow}$E), for, in many such cases, we're quantifying over at least some objects *which we don't have any access to*, and so about which—in the relevantly strong sense of 'about'—we're not in a position to grasp any content, and so about which we're not in a position to have any beliefs. Fortunately, there would seem to be an attitude of *accepting* more flexible than believing in that it allows for *accepting a content even if one does not grasp that content*, as evidenced by the fact that it would seem that quite a few people do accept the terms and conditions of many contracts, or the whole periodic table, or everything said by any past or future Pope etc. (I'll henceforth use 'accept' and its like to express such attitude; see [33] for some further discussion of the difference between acceptance and belief).

Thirdly, believing would also seem to be unsuitable for being *A*ing since *believing that the premises hold together would seem to be grounded in believing each premise individually*, whereas, in many cases in which it is plausible that we do actually use infinitary instances of ($T^{\Rightarrow}E$), *it is plausible that it is one's Aing that the premises hold together that grounds one's Aing each premise individually* (in this respect, *A*ing, contrary to believing, would belong to the same class of broadly intensional actions as buying, since it is possible directly to buy, say, all the mustard seeds together and only indirectly as a consequence of that buy each mustard seed individually—a similarity that should be expected given that *A*ing is in effect a sort of metaphorical buying!). Fortunately, accepting is an attitude more flexible than believing also in that it allows for *accepting that certain contents hold together to ground accepting each content individually*, as evidenced by the fact that it would seem that quite a few people do accept the terms and conditions of many contracts by directly accepting them all together, and only indirectly as a consequence of that accepting each of them individually, or accept the whole periodic table by directly accepting all its entries together, and only indirectly as a consequence of that accepting each of them individually, or accept everything said by any past or future Pope by directly accepting all those pronouncements together, and only indirectly as a consequence of that accepting each of them individually etc.

Fourthly, all one needs to do in order to use an infinitary instance of ($T^{\Rightarrow}E$) is to refer to the *plurality* of the infinitely many premises (under a suitable description) and accept that *they* hold *together*. The resources that one thereby employs are *more primitive* than those of full-blooded *universal quantification*, for, in accepting that the premises hold together, one need not employ the notion of *everythingness* and accept that *every* premise holds. Were it otherwise, the use of the relevant instance of ($T^{\Rightarrow}E$) would not really be one's point of entry to the relevant universal fact. On reflection, the resources one employs in using an infinitary instance of ($T^{\Rightarrow}E$) are *exactly the same* as those employed in a finitary, 2ary case in which one uses the rule of *adjunction* $\varphi, \psi \vdash \varphi \,\&\, \psi$. For, in that case too, all one needs to do in order to use an instance of the rule of adjunction is to refer to the *plurality* of the two premises (under a suitable description) and accept that *they* hold *together*. The resources that one thereby employs are *more primitive* than those of full-blooded *conjunction*, for, in accepting that the premises hold together, one need not employ the notion of *andness* and accept that one premise holds *and* the other premise holds. Were it otherwise, the use of the relevant instance of the rule of adjunction would not really be one's point of entry to the relevant conjunctive fact. Thus, in both cases, one employs exactly the same resources allowing one to refer to a plurality of premises and accept that they hold together—the only difference is that in the case of an infinitary instance of ($T^{\Rightarrow}E$) the premises referred to are infinitely many, while in the case of the rule of adjunction the premises referred to are 2.

Fifthly, although, in using an infinitary instance of ($T^{\Rightarrow}E$), as *per* the second last paragraph one need not directly accept of each premise that it holds, in those cases in which it is plausible that we do actually use infinitary instances of ($T^{\Rightarrow}E$) it is also plausible that, in a suitably idealised sense, one is indeed *disposed* directly to accept of each premise that it holds. This is so because, in those cases, for each premise one will be *in possession of a form of ground that one in principle knows how, and so is disposed, to apply to that premise* to produce a specific ground for the premise.[37] A finite mind may not capable of infinitely many *occurrent thoughts*, but it is capable of infinitely many *standing dispositions*—indeed, of as many dispositions as there are objects in the universe (consider, for every object in the universe, our disposition to accept that it is self-identical). I submit that it is these dispositions that provide the bases on which, by a process of synthesis, one can then accept that the premises of the relevant instance of ($T^{\Rightarrow}E$) hold together.

Something like this account strikes me as the only plausible account of how we come to accept, say, $\&_{\mathbf{ack}(1{,}000{,}000)}(0 = 0, 1 = 1, 2 = 2 \ldots, \mathbf{ack}(1{,}000{,}000) - 1 = \mathbf{ack}(1{,}000{,}000) - 1)$ by using ($\&_{\mathbf{ack}(1{,}000{,}000)}$-INTRO). But, once we've recognised that we need anyways something like the account I've sketched to rationalise that acceptance, it becomes open to us to deploy the same account also to rationalise our acceptance of (SUCC), and, more generally, to understand how, in spite of our finitude, we can actually use infinitary instances of ($T^{\Rightarrow}E$). And, if we do so, we'll have an account according to which *our reasoning about everythingness is uniform across all cases*, the only difference being that, because of our limitations, in some cases our use of the very same principle as we straightforwardly use in a case like the class example is perforce more roundabout.

21.6 Going Beyond

Let's close by identifying the *crux* of the matter, at least as far as **LW** and $\mathcal{N}_{\mathbf{LS}}$ are concerned. Reflect that, in both those theories, the failure of ($^{\Rightarrow}C_AE$) can be traced back to the failure of a certain very strong version of the *metarule* of *adjunction*[38]:

(MAD) If $\Gamma \vdash \varphi$ and $\Gamma \vdash \psi$ hold, $\Gamma \vdash \varphi \,\&\, \psi$ holds.

In the case of **LW**, both $\lambda \vdash_{\mathbf{LW}} \lambda$ holds (by reflexivity of logical consequence) and $\lambda \vdash_{\mathbf{LW}} \neg\lambda$ holds (by reflexivity of logical consequence and transparency), but $\lambda \vdash_{\mathbf{LW}} \lambda \,\&\, \neg\lambda$ does not hold (if it did, by a special version of *reductio ad absurdum* $\oslash \vdash_{\mathbf{LW}} \neg\lambda$ would hold).[39] In the case of $\mathcal{N}_{\mathbf{LS}}$, both $\mathcal{N}, \neg\beta_2 \vdash_{\mathbf{LS}} \beta_1$

[37] I conjecture that it is at this more *psychological* level of *dispositions*—rather than at the more *logical* level of *contents*—that, suitably reconfigured, some stuff about "*schematic reasoning*" might come in useful.

[38] On the contrary, the *rule* of adjunction mentioned in Sect. 21.5 is valid in both theories.

[39] $\oslash$ is the *empty multiset*.

holds (by $\mathcal{N}$'s including β_0 as well as $\beta_0 \supset \beta_1$, *modus ponens* and monotonicity of logical consequence) and $\mathcal{N}, \neg\beta_2 \vdash_{\mathbf{LS}} \neg\beta_2$ holds (by reflexivity and monotonicity of logical consequence), but $\mathcal{N}, \neg\beta_2 \vdash_{\mathbf{LS}} \beta_1 \,\&\, \neg\beta_2$ does not hold (if it did, since $\mathcal{N}$ includes $\beta_1 \supset \beta_2$, by another special version of *reductio ad absurdum* $\mathcal{N} \vdash_{\mathbf{LS}} \beta_2$ would hold). Thus, (MAD) (and hence ($^{\Rightarrow}$C$_{\mathrm{A}}$E)) has got to fail if anything like the theory of truth embodied in **LW** and the theory of baldness embodied in $\mathcal{N}_{\mathbf{LS}}$ are on the right track. These two arguments already offer an adequate *logical* explanation of the failure of (MAD), for the arguments show in detail precisely why (MAD) is logically uncotenable with other compelling principles (in particular, with the *exclusivity* and *exhaustivity* of Boolean negation and either *transparency* of truth or *tolerance* of baldness). Yet, can **LW** and $\mathcal{N}_{\mathbf{LS}}$ also offer a fuller, *philosophical* explanation of the failure of (MAD)?

I think they can. On both theories, reality is *radically fragmented*, in the sense that *there is no complete way things are*. In the case of the philosophical view underpinning **LW**, this is so because of the *dynamicity* of the states-of-affairs (SOAs) expressed by self-referring, truth-predicating sentences. For *transparency* gives rise to an endless sequence of *stages of truth evaluation*, where each stage leads to another stage *inconsistent* with it. For example, a stage at which λ is evaluated as not true leads, by transparency, to a stage at which λ is evaluated as true, and, because of the exclusivity of Boolean negation, any two such stages are inconsistent. Now, any putative complete way things are will have to contain, by completeness, either the SOA of λ being not true or the SOA of λ being true, but not both (since the two SOAs are inconsistent, and so anything containing them both will not be a way *things are*). Whichever it includes, the putative complete way things are will thus only correspond to a specific stage of truth evaluation, leading to another stage which is inconsistent with it, and so which includes a SOA that is not contained in the putative complete way things are: therefore, that way will not be a *complete* way things are. Thus, although both the SOA expressed by λ and the SOA expressed by $\neg\lambda$ belong to some stages of truth evaluation, there cannot be any single, overarching, maximally specific fact of which they are both part: in this sense, they are only parts of two essentially disjoint fragments of reality. And this fragmentation generalises to many other pairs of, on the one hand, the SOA expressed by a self-referring, truth-predicating sentence and, on the other hand, the SOA expressed by the sentence attributing truth to the original sentence (for example, the SOA expressed by a *Curry* sentence and the SOA expressed by the sentence attributing truth to that Curry sentence).[40]

Nevertheless, since both the SOA expressed by λ and the SOA expressed by $\neg\lambda$ belong to some stages of truth evaluation, it must be *possible* to *go* from the untruth of the Liar sentence to the truth of the Liar sentence. The *possibility* of this transition at the *metaphysical* level corresponds at the *logical* level to the *consistency* of theories the development of whose logical consequences represents some such

[40]Does this require *distinguishing* between the SOA expressed by φ and the SOA expressed by ' 'φ' is true' to an extent incompatible with a *deflationist* conception of truth? Probably.

transition, its beginning and its end. In the framework of **LW**, one particularly relevant such theory is λ, since, on the hand, $\lambda \vdash_{\mathbf{LW}} \lambda$ holds, and so the situation λ describes can be taken to be a situation in which the Liar sentence is not true (the beginning of the transition), while, on the other hand, $\lambda \vdash_{\mathbf{LW}} \neg\lambda$ holds, and so the situation λ describes can equally well be taken to be, on the contrary, a situation in which the Liar sentence is true (the end of the transition). As we now know that and why the SOAs expressed by λ and $\neg\lambda$ are part of two essentially disjoint fragments of reality, we can now understand why (MAD) fails for λ and for other consistent theories that can similarly swing [between two essentially disjoint fragments of reality] representing the transition from a SOA that is part of one to a SOA that is part of the other: in such cases, (MAD) would in effect force such consistent theories to put together those SOAs into a single, overarching, virtually maximally specific[41] and therefore impossible SOA.[42]

In the case of the philosophical view underpinning $\mathcal{N}_{\mathbf{LS}}$, reality is radically fragmented because of the *constancy* of the SOAs expressed by baldness-predicating sentences. For *tolerance* gives rise to a full sequence of *pairwise equivalent cases*. For example, either both 0 and 1 are such that a man with that number of hairs is bald or neither is; either both 1 and 2 are such that a man with that number of hairs is bald or neither is; either both 2 and 3 are such that a man with that number of hairs is bald or neither is etc., and, because of the exclusivity of Boolean negation, these equivalences are inconsistent with a man with 0 hairs being bald but a man with 1 hair not being bald, with a man with 1 hair being bald but a man with 2 hairs not being bald, with a man with 2 hairs being bald but a man with 3 hairs not being bald etc. Now, any putative complete way things are will have to contain, by completeness, both the SOA of a man with 0 hairs being bald and the SOA of a man with 999,999 hairs not being bald. Containing both, since, by completeness, the putative complete way things are will have to contain also a verdict about every case in between, the putative complete way things are will end up containing, for some i, both the SOA of a man with i hairs being bald and the SOA of a man with $i + 1$ hairs not being bald, which contradicts tolerance: therefore, that way will not be a complete way *things are*. Thus, although both the SOA expressed by β_0 and the SOA expressed by $\neg\beta_{999,999}$ obtain, there cannot be any single, overarching, maximally specific fact of which they are both part: they are only part of two different, essentially disjoint fragments of reality. And this fragmentation generalises to many other pairs of, on the one hand, the SOA a

[41] Although the SOA of the Liar sentence not being true and the SOA of the Liar sentence being true obviously do not constitute a maximally specific SOA, they *virtually* do so, since, for our purposes, there would be no further bar to adding to them all the other relevant SOAs (like e.g. the SOA of 'Snow is white' being true).

[42] A loose analogy that some might find helpful: consider a (MAD)ish principle for *pictorial representation* to the effect that, if what a picture represents is an F and what it represents is a G, what it represents is an F and G. Whatever its other merits, such principle dramatically fails for pictures generating *Gestalt*-switches.

positive case consists in and, on the other hand, the SOA a negative case consists in (for example, the SOA of 999,999 grains making a *heap* and the SOA of 0 grains not making a heap).

Nevertheless, since both the SOA expressed by β_0 and the SOA expressed by $\neg\beta_{999,999}$ hold, it must be *possible* to *go* from the baldness of a man to the non-baldness of a man. The *possibility* of this transition at the *metaphysical* level corresponds at the *logical* level to the *consistency* of theories the development of whose logical consequences represents some such transition, its beginning and its end. In the framework of $\mathcal{N}_{\mathbf{LS}}$, one particularly relevant such theory is $\mathcal{N}, \neg\beta_2$, since, on the hand, $\mathcal{N}, \neg\beta_2 \vdash_{\mathbf{LS}} \beta_1$ holds, and so the situation $\mathcal{N}, \neg\beta_2$ describes can be taken to be a situation in which a man with 1 hair is bald (the beginning of the transition), while, on the other hand, $\mathcal{N}, \neg\beta_2 \vdash_{\mathbf{LS}} \neg\beta_2$ holds, and so the situation $\mathcal{N}, \neg\beta_2$ describes can equally well be taken to be, on the contrary, a situation in which a man with 2 hairs is not bald (the end of the transition).[43] As we now know that and why the SOAs expressed by β_0 and $\neg\beta_{999,999}$ are part of two essentially disjoint fragments of reality, we can now understand why (MAD) fails for $\mathcal{N}, \neg\beta_2$ and for other consistent theories that can similarly swing [between two essentially disjoint fragments of reality] representing the transition from a SOA that is part of one to a SOA that is part of the other: in such cases, (MAD) would in effect force such consistent theories to put together those SOAs into a single, overarching, virtually maximally specific[44] and therefore impossible SOA.

Let's call theories for which (MAD) fails '*non-adjunctive*'. A non-adjunctive theory has, among other things, an interesting property: very plausibly, for some sentences φ and ψ (those for which the theory is non-adjunctive), the theory entails *either* φ *or* ψ*, but not both*. This is so since, by a very plausible connection

[43]It is obviously not the case that a man with 2 hairs is not bald, but the general point should be clear: given a theory of baldness $\mathcal{N}'$ *realistically stronger* than $\mathcal{N}$ (in all its three conjuncts), and assuming *classical logic in the metatheory*, a familiar Sorites-style reasoning yields that there is a least number i such that $\mathcal{N}' \vdash \beta_i$ does not hold, and what I say in text about $\mathcal{N}$ and β_2 holds for $\mathcal{N}'$ and β_i just as well. To avoid a possible misunderstanding, I'm emphatically *not* saying that we should accept $\mathcal{N}', \neg\beta_i$. Given the conception of the normativity of logical consequence congenial to **LS** mentioned in Footnote 14, assuming that we accept $\mathcal{N}'$ as our "basic theory" we should accept β_{i-1}, but it is not the case that we should accept β_i. In fact, plausibly, we should [not accept β_i] (since $\mathcal{N}' \vdash \beta_i$ does not hold). And, equally plausibly, we should [not accept $\neg\beta_i$] either (since what we would accept would then pin down a counterexample to tolerance). But, by exhaustivity of Boolean negation, we may not accept β_i ($\mathcal{N}' \vdash \beta_i$ does not hold) *only if* $\mathcal{N}', \neg\beta_i$ *is consistent*. *Too strong* as $\mathcal{N}', \neg\beta_i$ may be for us to accept it as a theory of the transition from the *baldness* of a man to the *non-baldness* of a man, under the assumptions made explicit in this fn it is exactly *its* consistency rather than that of some *weaker* theory that explains how, in spite of the constancy enforced by tolerance, it is possible to go from the $\mathcal{N}'$*-necessity* of the baldness of a man to the *non-*$\mathcal{N}'$*-necessity* of the baldness of a man (and so explains how *we* can nevertheless go from *accepting* that a man is bald to *not accepting* that a man is bald).

[44]Although the SOA of a man with 1 hair being bald and the SOA of a man with 2 hairs not being bald obviously do not constitute a maximally specific SOA, they *virtually* do so, since, for our purposes, there would be no further bar to adding to them all the other relevant SOAs (like e.g. the SOA of a man with 0 hairs being bald).

between bothness and conjunction, a theory Γ entails both φ and ψ iff Γ entails $\varphi \,\&\, \psi$. Thus, if Γ is non-adjunctive for φ and ψ, Γ entails φ and Γ entails ψ, Γ entails either φ or ψ but Γ does not entail both φ and ψ. With respect to these 'each'/'either'/'not both'-features, a non-adjunctive theory behaves very much like the believer in the preface paradox and the counterfactual supposition in the party paradox of Sect. 21.4 (as I've also argued in Sect. 21.4, the permission in free-choice permission lacks the 'each'-feature).

Focussing on non-adjunctive theories like **LW** and $\mathcal{N}_{\mathbf{LS}}$, from the point of view of *what these theories represent* such 'either'/'not both'-features make sense as the theories represent transitions between SOAs s_0 and s_1 that are parts of two essentially disjoint fragments of reality. It is in the nature of such transitions to constitute a situation in which either s_0 or s_1 obtains (since, being a transition, the situation is tethered neither to s_0 nor to s_1, and so it can be taken to be such that s_0 does not hold in it, and hence such that s_1 holds in it, or can be taken to be such that s_1 does not hold in it, and hence such that s_0 holds in it) but not both (since s_0 and s_1 are part of two essentially disjoint fragments of reality).

From the point of view of *reasoning in such theories*, this opens up unexpected areas of *freedom*. To begin with, recall the familiar understanding of the "*hardness of the logical must*", according to which, if a certain theory Γ entails a certain consequence φ, under the assumption of Γ, if queried about φ etc., one *must* accept φ. But now suppose that Γ is non-adjunctive for φ and ψ. Then it is actually not the case that under the assumption of Γ, if queried about φ etc., one *must* accept φ—one *may* accept ψ *instead*, which would arguably exempt—indeed, prevent—one from accepting φ (since, in the relevant cases in which (MAD) fails in **LW** and $\mathcal{N}_{\mathbf{LS}}$, φ is inconsistent with ψ). Analogously, it is actually not the case that, under the assumption of Γ, if queried about ψ etc., one *must* accept ψ—one *may* accept φ *instead*, which would arguably exempt—indeed, prevent—one from accepting ψ (for the same reason as above). Thus, under the assumption of Γ, if queried about φ and ψ etc., one must in effect *decide* between them. And, since both options conform equally well to the facts about logical consequence, in the relevant respects such decision will be *arbitrary* (pretty much like the decision of how, roughly, to instantiate the relevant anythingness claims in the preface and party paradoxes as well as free-choice permission).[45] *Logical necessity* has been bent; a higher degree of *logical freedom*, springing from the fractures of reality, is revealed. Moreover, whichever of φ or ψ one decides for, *one will no longer be able to appeal to* Γ, since, if one could, one could then use it to infer the other conclusion too. Thus, sometimes, in inferring a conclusion one is *no longer bound* to the original assumptions: one has indeed inferred the conclusion *because of* the assumptions,

[45]Romance languages have typically been blessed with determiners expressing anythingness that suggestively reflect in their etymology the deep connection between anythingness and freedom: see for example Portuguese *qualquer*, Galician *calquera*, Spanish *cualquier*, Catalan *qualsevol*, Italian *qualsivoglia* and (likely) Romanian *orice* (as for Latin itself, see *quivis* and *quilibet*).

but the very acceptance of the conclusion *frees one from one's commitment* to the assumptions. In logical inference, one can not only be *free*, but also be *set free*.

Acknowledgements Earlier versions of the material in this paper have been presented in 2013 at the LOGOS Workshop *Substructural Approaches to Paradox* (University of Barcelona) and at the 7th Navarre Vagueness Workshop in Pamplona (University of Navarre), where Paul Égré gave a valuable response; in 2014, at the 3rd Colombian Congress on Logic, Epistemology and Philosophy of Science in Bogotá (University of the Andes/University of the Rosario) and at the PERSP Metaphysics Seminar (University of Barcelona). I'd like to thank all these audiences for very stimulating comments and discussions. Special thanks go to Pablo Cobreros, Anamaria Fălăuş, Nissim Francez, John Horden, Dan López de Sa, Pepe Martínez, Giovanni Merlo, Joe Moore, Julien Murzi, Sergi Oms, Francesco Paoli, Paloma Pérez-Ilzarbe, Dave Ripley, Sven Rosenkranz, Gonçalo Santos, Lionel Shapiro, Roy Sorensen, Luca Tranchini, Alan Weir and Dan Zeman. I'm also grateful to the editor Alessandro Torza for inviting me to contribute to this volume and for his support and patience throughout the process. At different stages during the writing of the paper, I've benefitted from the FP7 Marie Curie Intra-European Research Fellowship 301493 on *A Non-Contractive Theory of Naive Semantic Properties: Logical Developments and Metaphysical Foundations* (NTNSP) and from the FCT Research Fellowship IF/01202/2013 on *Tolerance and Instability: The Substructure of Cognitions, Transitions and Collections* (TI), as well as from partial funds from the project CONSOLIDER-INGENIO 2010 CSD2009-00056 of the Spanish Ministry of Science and Innovation on *Philosophy of Perspectival Thoughts and Facts* (PERSP), from the FP7 Marie Curie Initial Training Network 238128 on *Perspectival Thoughts and Facts* (PETAF), from the project FFI2011-25626 of the Spanish Ministry of Science and Innovation on *Reference, Self-Reference and Empirical Data* and from the project FFI2012-35026 of the Spanish Ministry of Economy and Competition on *The Makings of Truth: Nature, Extent, and Applications of Truthmaking*.

References

1. Ackermann, W. 1928. Zum Hilbertschen Aufbau der reellen Zahlen. *Mathematische Annalen* 99: 118–133.
2. Adams, E. 1970. Subjunctive and indicative conditionals. *Foundations of Language* 6: 89–94.
3. Anderson, A. 1956. *The formal analysis of normative systems*. New Haven: Yale University Press.
4. Armstrong, D. 1997. *A world of states of affairs*. Cambridge: Cambridge University Press.
5. Esenin-Vol'pin, A. 1970. The ultra-intuitionistic criticism and the antitraditional program for foundations of mathematics. In *Intuitionism and proof theory*, ed. A. Kino, J. Myhill, and R. Vesley, 3–45. Amsterdam: North-Holland.
6. Evans, G. 1978. Can there be vague objects? *Analysis* 38: 208.
7. Fiengo, R. 2007. *Asking questions*. Oxford: Oxford University Press.
8. Fine, K. 1983. A defence of arbitrary objects. *Proceedings of the Aristotelian Society Supplementary Volume* 57: 55–77.
9. Heidegger, M. 1929. *Was ist Metaphysik?* Bonn: Friedrich Cohen.
10. Kyburg, H. 1961. *Probability and the logic of rational belief*. Middletown: Wesleyan University Press.
11. Leonard, H. 1956. The logic of existence. *Philosophical Studies* 7: 49–64.
12. Makinson, D. 1965. The paradox of the preface. *Analysis* 25: 205–207.
13. McGee, V. 1985. A counterexample to modus ponens. *The Journal of Philosophy* 82: 462–471.
14. Paoli, F. 2005. The ambiguity of quantifiers. *Philosophical Studies* 124: 313–330.
15. Read, S. 1981. What is wrong with disjunctive syllogism? *Analysis* 41: 66–70.

16. Russell, B. 1903. *The principles of mathematics*. Cambridge: Cambridge University Press.
17. Vogel, J. 2000. Reliabilism leveled. *The Journal of Philosophy* 97: 602–623.
18. von Wright, G. 1951. Deontic logic. *Mind* 60: 1–15.
19. Weir, A. 2005. Naive truth and sophisticated logic. In *Deflationism and paradox*, ed. B. Armour-Garb and J.C. Beall, 218–249. Oxford: Oxford University Press.
20. Weir, A. 2010. *Truth through proof*. Oxford: Oxford University Press.
21. Wittgenstein, L. 1921. Logisch-philosophische Abhandlung. *Annalen der Naturphilosophie* 14: 185–262.
22. Wright, C. 1982. Strict finitism. *Synthese* 51: 203–282.
23. Zardini, E. 2008a. A model of tolerance. *Studia Logica* 90: 337–368.
24. Zardini, E. 2008b. *Living on the slippery slope. The nature, sources and logic of vagueness*. Ph.D. thesis, Department of Logic and Metaphysics, University of St. Andrews.
25. Zardini, E. 2009. Towards first-order tolerant logics. In *Philosophy, mathematics, linguistics: Aspects of interaction*, ed. O. Prozorov, 35–38. St. Petersburg: Russian Academy of Sciences Press.
26. Zardini, E. 2011. Truth without contra(di)ction. *The Review of Symbolic Logic* 4: 498–535.
27. Zardini, E. 2013a. It is not the case that [*P* and 'It is not the case that *P*' is true] nor is it the case that [*P* and '*P*' is not true]. *Thought* 1: 309–319.
28. Zardini, E. 2013b. Naive *modus ponens*. *Journal of Philosophical Logic* 42: 575–593.
29. Zardini, E. 2013c. Closed without boundaries. MS.
30. Zardini, E. 2014a. Confirming the less likely, discovering the unknown. Dogmatisms: Surd and doubly surd, natural, flat and sharp. In *Scepticism and perceptual justification*, ed. D. Dodd and E. Zardini, 33–70. Oxford: Oxford University Press.
31. Zardini, E. 2014b. Evans tolerated. In *Vague objects and vague identity*, ed. K. Akiba and A. Abasnezhad, 327–352. Berlin: Springer.
32. Zardini, E. 2014c. Naive truth and naive logical properties. *The Review of Symbolic Logic* 7: 351–384.
33. Zardini, E. 2014d, Forthcoming. The opacity of truth. *Topoi*.
34. Zardini, E. 2015. Breaking the chains. Following-from and transitivity. In *Foundations of logical consequence*, ed. C. Caret and O. Hjortland, 221–275. Oxford: Oxford University Press.
35. Zardini, E. 2015. Getting one for two, or the contractors' bad deal. Towards a unified solution to the semantic paradoxes. In *Unifying the philosophy of truth*, ed. T. Achourioti, K. Fujimoto, H. Galinon, and J. Martínez, 461–493. Berlin: Springer.
36. Zardini, E. Forthcoming. És la veritat una mentida? Perspectives sobre les paradoxes semàntiques. *Anuari de la Societat Catalana de Filosofia*.
37. Zardini, E. Forthcoming. First-order tolerant logics. *The Review of Symbolic Logic*.
38. Zardini, E. Forthcoming. Restriction by non-contraction. *Notre Dame Journal of Formal Logic*.