Logic, Epistemology, and the Unity of Science 23

Shahid Rahman
Giuseppe Primiero
Mathieu Marion *Editors*

The Realism-Antirealism Debate in the Age of Alternative Logics

Springer

The Realism-Antirealism Debate in the Age of Alternative Logics

LOGIC, EPISTEMOLOGY, AND THE UNITY OF SCIENCE

VOLUME 23

Logic, Epistemology, and the Unity of Science aims to reconsider the question of the unity of science in light of recent developments in logic. At present, no single logical, semantical or methodological framework dominates the philosophy of science. However, the editors of this series believe that formal techniques like, for example, independence friendly logic, dialogical logics, multimodal logics, game theoretic semantics and linear logics, have the potential to cast new light on basic issues in the discussion of the unity of science.

This series provides a venue where philosophers and logicians can apply specific technical insights to fundamental philosophical problems. While the series is open to a wide variety of perspectives, including the study and analysis of argumentation and the critical discussion of the relationship between logic and the philosophy of science, the aim is to provide an integrated picture of the scientific enterprise in all its diversity.

For further volumes:
http://www.springer.com/series/6936

Shahid Rahman · Giuseppe Primiero ·
Mathieu Marion
Editors

The Realism-Antirealism Debate in the Age of Alternative Logics

Springer

Editors
Shahid Rahman
Université de Lille, UdL3
CNRS, UMR 8163: STL
F-59653 Villeneuve d'Ascq
France
shahid.rahman@univ-lille3.fr

Dr. Giuseppe Primiero
FWO, Ghent University
Blandijnberg 2
Ghent
Belgium

IEG
Oxford University
Oxford
UK
giuseppe.primiero@ugent.be

Mathieu Marion
Université du Québec à Montréal
Département de philosophie
Succursale Centre-Ville 8888 C.P.
H3C 3P8 Montreal
QC, Canada
marion.mathieu@uqam.ca

ISBN 978-94-007-1922-4 e-ISBN 978-94-007-1923-1
DOI 10.1007/978-94-007-1923-1
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2011937177

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the memory of Paul Gochet

Preface

In the preface to the first volume of LEUS, the editors of the series pointed out, within the context of the failure of the positivist project of the unity of science, the difference between science as a body of knowledge and science as process by which knowledge is achieved. In fact, the editors suggested that a ban on the logical analysis of science as a dynamic process, which in traditional philosophy was overtaken by ‘gnoseology’, produced a gap between sciences and logic (including philosophy of science). In gnoseology the main notion was the one of judgement rather than that of proposition. Judgement delivered the epistemic aspect of logic, namely the relation between an (epistemic) agent and a proposition. This represented the basis of the Kantian approach to logic, which seemed to be in conflict with the post-Fregean approach where only relations between propositions are at stake and where the epistemic aspect is seen as outside logic.

As it happens quite often in philosophy, the echoes of the old traditions come back and point to the mistakes of the younger iconoclast movements. This is indeed the case in the relation between logic and knowledge where the inclusion or exclusion of the epistemic moment as linked with the concept of proposition provoked a heated debate since the 1960s. The epistemic approaches, which started to call themselves, following Michael Dummett, ‘antirealism’, found their formal argument in the mathematics of Brouwer and intuitionistic logic, while the others persisted with the formal background of the Frege-Tarski tradition, where Cantorian set theory is linked via model theory to classical logic. This picture is, however, incomplete. Already in the 1960s Jaakko Hintikka tried to join both traditions by means of what is now known as ‘explicit epistemic logic’, where the epistemic content is introduced into the object language as an operator which yield propositions from propositions rather than as metalogical constraint on the notion of inference. The debate had thus three players: classical logicians, intuitionistic logicians (implicit epistemic logic) and epistemic logicians (explicit epistemic logic), though the mainstream continued to think that the discussion reduces to the discussion between classical and intuitionistic logic.

The editors of the present volume think that in these days and age of Alternative Logics, where manifold developments in logic happen in a breathtaking pace, this debate should be revisited. In fact, collaborators to this volume took happily this challenge and responded with new perspectives on the debate from both the explicit and the implicit point of view, challenging it from the newly arisen perspectives in logic. This volume aims therefore at presenting standard issues of the realism-antirealism debate in a new light, shed from the point of view of different philosophical perspectives. It is therefore appropriate that we open with Patrick Allo's contribution, which analyses the meaning of ambiguous connectives (and in particular of disjunction) from a logical pluralistic viewpoint, in which content is explained in terms of informativeness. Logical pluralism can be understood as the larger conceptual umbrella under which one finds today many different understandings of the realism-antirealism debate. This certainly still grows on Dummett's arguments against truth-conditional semantics, which Neil Kennedy reconstructs and critically analyses; in general, it refers to well known forms of semantic antirealism, which Sanford Shieh characterizes by means of the distinction between epistemically-based and conceptually based ones, and it still triggers today huge debates such as the one on Moorean validities, that Jon Cogburn reconstructs in view of the different old and new interpretations. But antirealism today profits of the influences of many different backgrounds: this is the case for example of Martin-Löf's type theory, which is conceptually and historically located within the larger frame of theories of truth and judgement in Göran Sundholm's contribution, and which meets for the first time belief revision dynamics in Giuseppe Primiero's paper. Departure from classical principles of reasoning is therefore possible in different forms, and whereas Denis Bonnay and Mikaël Cozic place the justification of radical forms of anti-realism in the context of the (to them still unjustified) shift to linear logic, Joseph Vidal-Rosset suggests a larger philosophical frame for the understanding of radical antirealism. Many are therefore the new branches of logic that are called upon in this volume to face non-classical issues raising from an antirealistic perspective: this is the case of modal logic in the interpretations by Reinhard Kahle and Elia Zardini; the anti-realistic inspired defence of realist mathematics by Greg Restall, where (implicit) antirealism is understood as a means to defend logical pluralism: the extension to Paraconsistency, via a defence of a suitable negation connective in the Kripke-Hintikka reconstruction of intuitionistic logic, suggested by Graham Priest, which virtually dialogues with the dialogical interpretation of the same connective given by Shahid Rahman. The relation of antirealism to dialogical logic and game semantics appears also in Mathieu Marion's contribution, where it is considered how to make Dummett's Manifestation Argument work within this new programme, and it is argued that a derived Thesis fits (with appropriate reformulations) within game semantics. Stephen Read, takes the proof theoretical approach to logic of the antirealists to challenge the epistemic constraints of the intuitionists and Francesca Poggiolesi analyses properties of anti-realistic definitions starting from the classical requirements imposed by Lesniewski. Truth is of course always an open field for new interpretations: Alexandre Billon considers the notion of assessment-sensitive truth to provide solutions to semantic paradoxes and Maria Frapolli presents a prosentential account of truth showing that our comprehension of truth and the use we make of truth expressions are strictly independent of our views about the relation between mind and world.

Some years have passed from the initial proposal to collect a number of contributions from scholars that are reconsidering the realism-antirealism debate from new philosophical, logical and metaphysical perspectives. This has led to different line-ups of both authors and editors. What we hope to have achieved through this process is to have selected significant contributions on the different aspects of research on anti-realism done today at academic level, a representative body of work that can be of reference and inspiration for further advancements in this field.

Montreal, Canada — Mathieu Marion
Ghent, Belgium — Giuseppe Primiero
Lille, France — Shahid Rahman
December 2010

Acknowledgements

This volume originates from the Workshop *(Anti-)Réalismes. Logique et Métaphysique*, which took place in Nancy (France), from June 28 to July 2, 2006. For that reason, we would like first to express our gratitude to the organizing committee formed by the *Laboratoire d'Histoire des Sciences et de Philosophie—Archives Henri Poincaré*, UMR 7117 CNRS, Nancy-Université and the laboratory UMR 8163: *Savoirs, Textes, Langage*, Université de Lille 3, with special thanks to Gerhard Heinzmann (Nancy 2), Laurent Rollet, Manuel Rebuschi (Nancy 2), Christian Berner (Lille3). The editors would also like to thank Laurent Keiff (UMR:STL), who started this project and later handed it to us. We also would like to mention here the work of the various readers and referees that helped us selecting and improving the contributions to this volume; their silent help is very much appreciated.

Contents

Contributors

Patrick Allo Postdoctoral Fellow, Fund for Scientific Research—Flanders (Belgium), Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium; IEG, Oxford University, Oxford, UK, patrick.allo@vub.ac.be

Alexandre Billon Université Lille Nord de France, F-59000 Lille, France; UdL3, STL, CNRS, UMR 8163, F-59653 Villeneuve d'Ascq, France, alexandre.billon@m4x.org

Denis Bonnay Université Paris Ouest-Nanterre, Ireph & IHPST, Paris, France, denis.bonnay@ens.fr

Jon Cogburn Louisiana State University Baton Rouge, Baton Rouge, LA, USA, joncogburn@yahoo.com

Mikaël Cozic Université Val de Marne & IHPST, Paris, France, Mikael.Cozic@ens.fr

María J. Frápolli Departamento de Filosofía I, Universidad de Granada, Granada, Spain, frapolli@ugr.es

Reinhard Kahle CENTRIA and DM, FCT, Universidade Nova de Lisboa, P-2829-516 Caparica, Portugal, kahle@mat.uc.pt

Neil Kennedy Université du Québec à Montréal, Montréal, QC, Canada; Université de Paris 1–IHPST, Paris, France, neil.patrick.kennedy@gmail.com

Mathieu Marion Canada Research Chair in the Philosophy of Logic and Mathematics, Université du Québec à Montréal, Montréal, QC, Canada, marion.mathieu@uqam.ca

Francesca Poggiolesi CLWF, Vrije Universiteit Brussel, Brussels, Belgium, poggiolesi@gmail.com

Graham Priest Departments of Philosophy, University of Melbourne, Melbourne, VIC, Australia; University of St Andrews, St Andrews, UK; The Graduate Center, City University of New York, New York, NY, USA, g.priest@unimelb.edu.au

Giuseppe Primiero FWO, Ghent University, Blandijnberg 2, Ghent, Belgium; IEG, Oxford University, Oxford, UK, Giuseppe.Primiero@UGent.be

Shahid Rahman Université de Lille, UdL3, CNRS, UMR 8163: STL, F-59653 Villeneuve d'Ascq, France, shahid.rahman@univ-lille3.fr

Stephen Read University of St Andrews, St Andrews, Scotland, UK, slr@st-andrews.ac.uk

Greg Restall School of Historical and Philosophical Studies, Melbourne, VIC, Australia, greg@consequently.org

Sanford Shieh Department of Philosophy, Wesleyan University, Middletown, CT 06459, USA, sanford.shieh@wesleyan.edu

Göran Sundholm Institute for Philosophy, Leiden University, Leiden, The Netherlands, goran.sundholm@gmail.com

Joseph Vidal-Rosset Archives Poincaré, UMR 7117, Nancy, France, joseph.vidal-rosset@univ-nancy2.fr

Elia Zardini Department of Philosophy, Northern Institute of Philosophy, University of Aberdeen, Aberdeen, UK, elia.zardini@abdn.ac.uk

Chapter 1
On When a Disjunction Is Informative

Ambiguous Connectives and a Realist Commitment to Pluralism

Patrick Allo

1.1 Pluralism About Consequence and Content

Following a suggestion of Beall and Restall's, it is our aim in this paper to exploit logical pluralism as a means to recognise a distinct, but not unrelated kind of pluralism. We refer to the latter as informational pluralism. Our starting point consists in two distinct arguments in favour of logical pluralism.

First, we have the argument based on an ambiguity pointed out by Beall and Restall [8, 9] in what they call the Generalised Tarski Thesis (GTT).

> (GTT) A conclusion A follows from premises Σ iff any *case* in which each premise in Σ is true is also a *case* in which A is true. [9, p. 29]

Their argument for logical pluralism specifically rests on an ambiguity in the use of the notion of a *case* within any fairly standard description of logical consequence as *truth-preservation*. Avoiding the traditional, but not always explicitly named restriction to complete and consistent cases, the *GTT* lays the ground for more than one logical system.

A second argument is a modification of an early objection to relevant logic due to Hanson [15], and relies on an alternative account of logical consequence as *content-nonexpansion*.

> (CN) An argument is valid just in case the content of the conclusion does not exceed the combined contents of the premises.

Whereas the first argument clearly and intendedly aims at a pluralist conclusion, the second one was originally presented with a monist interpretation in mind. Showing that a coherent account of content could be given such that *CN* exactly yields the classical notion of consequence, Hanson wishes to defuse the traditional relevantist's claim that logical consequence as truth-preservation ought to be extended

P. Allo (✉)
Postdoctoral Fellow, Fund for Scientific Research—Flanders (Belgium), Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

IEG, Oxford University, Oxford, UK
e-mail: patrick.allo@vub.ac.be

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_1,

with content-preservation. As a consequence, his provision of a classically minded notion of content suffices to show that if consequence needs both truth- and content-preservation, then classical logic is just fine and the mainstream relevantist incentive for logical revision readily vanishes.

The first claim to be defended in this paper is that Hanson's argument only works as an objection to, on the one hand, the relevantist monist, and, on the other hand, those relevant logicians who adhere to Anderson and Belnap's traditional (and admittedly rather vague) account of relevance. It will thus be argued that *CN* provides an equally compelling reason to be a logical pluralist as the *GTT* does in Beall and Restall's exposition. To do so we adopt *content* and *content-containment* as elementary notions, and argue for the thesis of informational pluralism using a suitably generalised version of the *Inverse Relationship Principle* [6].[1]

(GIRP) The informational content of a piece of information is given by the set of cases it excludes.

Jointly, *GIRP* and *GTT* provide all we need to sketch the analogy between Beall and Restall's original argument, and the one we wish to defend here. An early version of informational pluralism was introduced in [1] and further elaborated in [2]. Continuing along the lines set out in those papers, we specifically focus on issues related to the ambiguity of the classical disjunction, and thereby revisit some famous controversies that have been around since the systematic development of relevant logic by Anderson and Belnap. These are, to name but a few, the notion of relevance itself, the alleged opacity of the Routley-Meyer semantics, and the distinction between so-called intensional and extensional connectives.

The next two sections disambiguate the notion of content and rehearse the basics of its pluralist interpretation. In Section 1.4 an informal model that interprets the formal semantics for a weak relevant logic is given. Finally, this model is used to evaluate some objections against substructural logics and their pluralist interpretation.

1.2 Situated and Worldly Content

Adopting a strategy similar to Beall and Restall's, a pluralist individuation of the informational content of a message is obtained by varying among suitable precisifications of the notion of a *case* within the *GIRP*. Rather straightforwardly, this yields an account of classical, relevant, or intuitionistic content, depending on whether we let cases be possible worlds, situations, or constructions. Bypassing the underlying logical systems involved, the corresponding accounts of content can be derived directly on the basis of the modal space used in the respective frame-semantics. For reasons of focus, we choose to leave the intuitionist option aside, and consider its classically and relevantly inspired interpretations only.

[1] This accounts differs from Hanson's own, since he identifies the content of A with the consequence-set of A.

Technical Intermezzo

Spelling out the relevant interpretation of consequence and informational content requires a basic insight in the model-theory for relevant logic. In this technical intermezzo we list the definitions for the formal concepts we will need later on.

Definition 1 (Routley-Meyer Frame) A Routley-Meyer frame is a structure $(S, Log, \sqsubseteq_{\_}, *)$, where S is a set of situations, Log the subset of logical situations in S, $\sqsubseteq_{\_}$ a ternary relation on S, and $*$ a unary relation on S. A partial order $\sqsubseteq$ on S is defined as $s_1 \sqsubseteq s_2 := \exists c \in Log \;\&\; s_1 \sqsubseteq_c s_2$, and can be understood as a refinement-relation or information-ordering. Its core property is the persistence it enforces:

$$\text{if } s_1 \sqsubseteq s_2 \text{ and } s_1 \Vdash p \text{ then } s_2 \Vdash p$$

Further properties of the ternary relation $\sqsubseteq_{\_}$ are commutation (1.1), association (1.2), and downward closure on channels (1.3):

$$s_1 \sqsubseteq_c s_2 \Rightarrow c \sqsubseteq_{s_1} s_2 \tag{1.1}$$

$$\exists s \in Log \;\&\; s_1 \sqsubseteq_{c_1} s \sqsubseteq_{c_2} s_2 \Rightarrow \exists c \in Log \;\&\; s_1 \sqsubseteq_c s_2 \;\&\; c_1 \sqsubseteq_{c_2} c \tag{1.2}$$

$$s_1 \sqsubseteq_c s_2 \;\&\; c' \sqsubseteq c \Rightarrow s_1 \sqsubseteq_{c'} s_2 \tag{1.3}$$

While $*$ ought to comply with:

$$s_1 \sqsubseteq_c s_2 \Rightarrow s_2^* \sqsubseteq_c s_1^*$$
$$s = s^{**}$$

The ternary relation $\sqsubseteq_{\_}$ is asymmetric, transitive, but not reflexive; the reflexivity of $\sqsubseteq_c$ is equivalent to $c \sqsubseteq c^*$, which doesn't hold for arbitrary c, see Proposition 1. Furthermore, we introduce ; as a shorthand for channel-composition:

$$c = c_1 ; c_2 \;\text{ iff }\; c_1 \sqsubseteq_{c_2} c$$

Definition 2 (Satisfaction) An evaluation $\Vdash$ for a standard language including both lattice and group-theoretical connectives is given by the following clauses:

$$s \Vdash A \sqcap B \text{ iff } s \Vdash A \text{ and } s \Vdash B$$
$$s \Vdash A \sqcup B \text{ iff } s \Vdash A \text{ or } s \Vdash B$$
$$s \Vdash A \rightsquigarrow B \text{ iff } s^* \nVdash a \text{ or } s \Vdash B$$

$$s \Vdash \sim A \text{ iff } s^* \nVdash A$$

$$s \Vdash A \rightarrow B \text{ iff } \forall s_1, s_2 \in S \text{ where } s_1 \sqsubseteq_s s_2,\ s_1 \Vdash A \Rightarrow s_2 \Vdash B$$
$$s \Vdash A \oplus B \text{ iff } \forall s_1, s_2 \in S \text{ where } s_1 \sqsubseteq_s s_2,\ s_1^* \nVdash A \Rightarrow s_2 \Vdash B$$
$$s \Vdash A \otimes B \text{ iff } \exists s_1, s_2 \in S \text{ where } s_1 \sqsubseteq_{s_2} s,\ s_1 \Vdash A \,\&\, s_2 \Vdash B$$

Proposition 1 (Compatibility and Consistency) *The existence of inconsistent or impossible situations is warranted by the failure of* $s \sqsubseteq s^*$*:*

$$s \Vdash p \sqcap \sim p \Rightarrow s \Vdash p \,\&\, s^* \nVdash p \text{ hence } s \not\sqsubseteq s^*$$

More generally, $\sqsubseteq$ *and* $*$ *encode what it means for two situations* s_1 *and* s_2 *to be compatible as* $s_1 \sqsubseteq s_2^*$*, which is a symmetric relation.*[2]

$$\text{If } s_1 \sqsubseteq s_2^* \text{ then } s_2^{**} \sqsubseteq s_1^*$$
$$\text{hence } s_2 \sqsubseteq s_1^*$$

End of technical intermezzo.

Whilst the classical approach—originally elaborated in [5] in terms of state descriptions—quite expectedly identifies the content of A with the set of possible worlds which do not support A, a mere enlargement of the modal space suffices to obtain a relevantly inspired account of content. That is, using Routley-Meyer frames, we no longer evaluate formulae at possible worlds, but at cases which should neither be consistent nor complete. Following standard usage we call such cases situations; the consistent ones are referred to as *possible*, the inconsistent ones as *impossible*. We should furthermore remark that possible worlds represent a peculiar kind of situation, namely the complete and consistent ones. Even more than it is the case for Beall and Restall's logical pluralism, this last consideration is constitutive for our informational pluralism.

Note firstly that by using the frame-semantics one succeeds in identifying possible worlds as a well-defined class of situations.

$$s \Vdash A \text{ or } s \Vdash \sim A \text{ iff } s = s^*$$
$$\{s : s = s^*\} \subset S$$

This enables us to avoid any further reference to possible worlds as a primitive notion. Thus applying the intuitive characterisation of content expressed in the *GIRP* with respect to this enlarged set of logical points, we obtain two distinct content-individuations:

$$CONT_s(A) = S \setminus \{s : s \Vdash A\} \tag{1.4}$$
$$CONT_w(A) = \{s : s = s^*\} \setminus \{s : s = s^* \,\&\, s \Vdash A\} \tag{1.5}$$

[2] Actually, the symmetry of compatibility only requires $s \sqsubseteq s^{**}$ (which enforces double negation introduction) instead of $s = s^{**}$ (the semantic postulate for double negation equivalence).

In analogy with the interpretation of relevant implication as situated inference from [18], we propose to call (1.4) the individuation of the situated content of a message A. Likewise, we call (1.5) the worldly content of A. Showing the value of these two distinct notions, we need to argue that (a) they genuinely disagree on the content of some messages, and (b) they both correspond to something real.

Since both issues are extensively discussed in [2], we only briefly recapitulate the argument by considering a message of the form 'A `or` `not`-A'. Formally, the individuations given above give us little to doubt about. $CONT_w(A \texttt{ or not-}A) = \emptyset$, whereas $CONT_s(A \texttt{ or not-}A)$ encompasses the non-empty set of all situations which do not decide A.[3] Despite having obtained a genuine disagreement using both formal characterisations of content, we still need to explain what it means for one and the same message to be assigned a non-empty situated content, but at the same time also have an empty worldly content. Put simply, we hold that each evaluation refers to a distinct feature: [W] because worlds are complete, the sender of such a message could have been in any possible world (the empty worldly content), [S] but because situations can be incomplete, she could not have been in any situation in any such world (the non-empty situated content). Only an A-deciding situation could have provided the required (factual or explicit) evidence for a truthful assertion; a requirement trivially satisfied by possible worlds or worldly situations, but not by arbitrary situations.

Crucially, one should keep in mind that the existence of a worldly perspective is a peculiar issue. Not only can it fail to exist w.r.t. an arbitrary situation (i.e. $s \sqsubseteq s^*$ does not hold for all $s \in S$), but its existence is not a persistent property either. This very feature is customarily referred to as the existence of so-called impossible situations, situations that are formally inconsistent and hence have no refinement that is a possible world. Without discussion, this is the single most controversial aspect of the modal space we use. Yet, we shall not dwell upon this issue, but just mention a few points. Firstly, on the formal level impossible situations do not pose any problem. Since their advent, many interpretations of impossible situations have been given, but none of them shall ultimately concern us here (e.g. the essays in [26], or more recently [34]). The only explanation of impossible situations we need to endorse is fairly minimal. It solely depends on the fact that no information A that is supported by a situation s can convey the information that s has no impossible refinements. As a special case of this, we equally have that no A supported by s actually conveys the information that s itself is possible. In short, even though inconsistent situations are plainly impossible, no message can explicitly convey the information a receiver requires to conclude that the sender of that message does not find herself in an inconsistent situation. Though impossible situations can be ruled out as a matter of logic alone, no message can explicitly express the need to do so.

[3] Note that the situated content of A ($\{s : s \nVdash A\}$) does not express a proposition, and hence is not a persistent kind of content. Namely, it does not hold that if $s \nVdash A$, then $s' \nVdash A$ for all $s \sqsubseteq s'$. Such issues regarding the non-persistence of properties should be kept in mind.

1.3 Factual and Constraining Content

With the standard relevantist recapture of the disjunctive syllogism in mind, the analysis provided in the previous section could be questioned on several levels. Before answering such objections, it is better to rehearse the basic aspects of the controversy surrounding the classical inference-rule of disjunctive syllogism. Several doubts regarding the status of DS find their origin in Lewis' independent argument for *Ex Falso Quodlibet*, an argument-form rightly considered the basis of the most problematic paradox of material and strict implication.

$$\dfrac{\dfrac{A}{A \texttt{ or } B}(\texttt{or } I) \qquad \sim A}{B}(DS^{\texttt{or}})$$

Rejecting this argument in view of the lack of *relevance* of the premises for the conclusion, the mainstream relevantist answer to *disjunctive syllogism* has consisted in the rejection of the classical rule of $DS^{\texttt{or}}$. However, when confronted with the need to explain the general usefulness and (apparent) correctness of the very same inference-rule, relevant logicians generally provide a more fine-grained account of its rejection [28].[4] Namely, the invalid rule is to infer B from $A \sqcup B$ and $\sim A$, whereas its valid version is formalised in the rule $(DS^{\oplus})$ given below. To the contrary, the classical rule of *Addition* ($\texttt{or } I$) cannot be used to obtain $A \oplus B$ from either A or B, but is only acceptable as $(\sqcup I)$.

$$\dfrac{A \oplus B, \sim A}{A}(DS^{\oplus}) \quad \dfrac{A}{A \sqcup B}(\sqcup I) \quad \dfrac{B}{A \sqcup B}(\sqcup I)$$

Put this way, it is not hard to see why to the relevantist the classical *Ex Falso* rests on a spurious equivocation contained within the classical disjunction. Following the terminology used in [24], we call $\oplus$, and its dual $\otimes$ the group-theoretical disjunction and conjunction, whereas $\sqcup$ and $\sqcap$ are designated as lattice connectives. To avoid confusion, we stick to '$\texttt{or}$' and '$\texttt{and}$' when referring to ambiguous connectives.

The problem that this way out poses for a pluralist interpretation of informational content is threefold. Namely, (a) if correct, it renders the pluralist account of the content of a message '$A \texttt{ or not-}A$' incomplete, (b) it threatens the need for classical logic and hence also the need for a notion of worldly content, and (c) if only an ad hoc solution, the problem of explaining away the actual uses of DS (when there is no worldly perspective) remains. For starters, we do acknowledge that in order to provide an exhaustive account of informational content, the notions of situated and worldly content given above do not suffice as an explanation of the content of a message of the form '$A \texttt{ or } B$'. They need to be extended with

[4] A second, more complicated, aspect of the relevant recapture of DS, namely the admissibility of rule γ, which states that from $\vdash A \sqcup B$ and $\vdash \sim A$ we may derive that $\vdash B$, is left aside [3, §25].

a content-individuation that incorporates the frame-interpretation of $A \oplus B$. This provides an answer for the concern expressed under (a); solutions for the objections mentioned under (b) and (c) follow in Section 1.5.

In order to understand the content conveyed by (a message of the form) $A \oplus B$, two distinct perspectives can be adopted. First, reading of the relevant satisfaction-clauses we notice that while $s \Vdash A \sqcup B$ conveys information about the truth of A and B at s, it turns out that $s \Vdash A \oplus B$ remains silent on these matters. In other words, $\sqcup$ is extensional whilst $\oplus$ is intensional. This suggests that an all too simple interpretation of the *GIRP* with respect to messages correctly disambiguated using group-theoretical connectives might very well be fallacious. Indeed, it is an often made remark that the connective $\oplus$ primarily possesses inferential force. Not only does it satisfy a version of DS (which accounts for its inferential force), but the failure of addition points to the other side of this issue, namely its lack of factual content.

The solution we propose is to treat the content of $\sim A \oplus B$ as the conditional content of B, given A. A move warranted by (i) the equivalence of $A \rightarrow B$ and $\sim A \oplus B$, and (ii) the assumption that identifying the content of a conditional with the conditional content of the consequent given the antecedent is the right attitude to capture the inferential content of a message using $\oplus$ as its main connective.

Implementing this insight, we first need to recall that traditionally (worldly) conditional content is defined in terms of content simpliciter, see (1.6). As a corollary, we also note that when applied to worldly content, this way of defining conditional content does not discriminate between the factual and the inferential content of a message. The collapse given in (1.7) is inherent to the notion of worldly content, and plainly mimics the definition of material implication as $\sim A \vee B$.

$$CONT_w(B \mid A) = CONT_w(A \,\&\, B) \setminus CONT_w(A) \tag{1.6}$$

$$CONT_w(B \mid A) = CONT_w(\sim A \vee B) \tag{1.7}$$

However a similar equivalence between the content of B on the condition that A, and the content of $\sim A \oplus B$ is clearly intended, one must be cautious not to interpret the antecedent or condition as conveying factual content itself. In order to avoid this confusion, we first define the set of situations that are accessible *via* a channel c from a situation which carries the information that A, in short: c-accessible A-situations (see [19]), that shall serve as an individuation of the content conveyed by the antecedent as:

$$\begin{aligned} A_c^{\sqsubseteq} &:= \{s_i : \exists s(s \Vdash A \,\&\, s \sqsubseteq_c s_i)\} \\ &:= \{s_i : \exists s(s \Vdash A \,\&\, \forall B(c \Vdash B \Rightarrow s_i \Vdash A \otimes B))\} \end{aligned} \tag{1.8}$$

This can subsequently serve to give a content-individuation of B given A that does not collapse into the set of all situations that support A and $\sim B$. To see why, just consider that c-accessible A-situations need not carry the information that A even though they carry the information that $A \otimes B$ for any B that holds at c. By giving an intensional or group-theoretical reformulation of (1.6), it bypasses the factual interpretation of conditional content, and gives it a purely constraining interpretation instead.

$$\begin{aligned} \mathrm{CONT}_c(B \mid A) &= A_c^{\sqsubseteq} \setminus (A_c^{\sqsubseteq} \cap \{s : s \Vdash B\}) \\ &= A_c^{\sqsubseteq} \setminus \{s : s \Vdash A \otimes B\} \end{aligned} \tag{1.9}$$

Informally, the situated constraining content of B, given A is best understood as the proportion of c-accessible (or informationally linked) A-situations that do not eventually support B. Importantly, it can be shown that (1.9) individuates a proposition.

This leaves us with three distinct accounts of content: [W] standard worldly content, [S] situated factual content, and [C] inferential or situated constraining content. A few features of each of these need to be highlighted. In the first place we must conclude that worldly content accounts for both the inferential or constraining content as well as for the factual content of a message. Secondly, it can be noted that the strict distinction between situated factual content (which is extensional) and situated constraining content (which is intensional) is perfectly mirrored by the distinction between the lattice and the group-theoretical connectives. Finally, the fact that classical logic cannot discriminate between these two classes of connectives provides a sufficient ground to claim that from a worldly perspective the constraining content of a message *just is* its factual content. It is crucial to our pluralist enterprise that this latter fact is not dismissed as a spurious equivocation, but rather treated as a natural restriction on the discriminatory power that is inherent to the worldly perspective on content.

Before providing a model in which each of these notions receives a plausible interpretation, it can already be noted that all three [W], [S], and [C] provide a specific instance of a general insight into the nature of content and informativeness. Namely, the *dual to logic* intuition:

(DTL) X is informative only to the extent that X lies outside the scope of logical consequence alone.

It can thus easily be checked that [W] is dual to classical logic, [S] to relevant tautologies, and [C], provided it is only evaluated with respect to channels that are logical (i.e. $c \in Log$), is dual to relevant consequence.

1.4 Modelling Content

A simple model that allows one to interpret a formal model of informational content involves at least: (i) a sender, (ii) a receiver, and (iii) an event observed by the sender. What we evaluate in such a model is in the first place the content a message (truthfully) conveys to the receiver about an event the sender observes. As before [2], we choose a simple game as the event reported upon. A standard game-tree is used to establish what counts as the sender's evidential situation.

A game of *tic-tac-toe* is used as an actual example, and formalised in the following manner. Atomic formulae referring to a move in the game are of the form $P_i a_j$ where P_i is a position on a 3×3-board, and a_j refers to the turn in the game (such that a_1 is the first move, etc.).

$$\begin{aligned} \mathsf{Position} &= \{P_1, P_2, \ldots, P_9\} \\ \mathsf{Turn} &= \{a_1, a_2, \ldots a_9\} \\ \mathsf{Atom} &= \{P_i a_j \mid P_i \in \mathsf{Pos} \ \& \ a_j \in \mathsf{Turn}\} \end{aligned}$$

A basic description of an evidential situation s_i or sequence of board-configurations is a finite conjunction of atomic formulae. For instance, $s_3 \Vdash P_1 a_1 \sqcap P_5 a_2 \sqcap P_7 a_3$ describes the board resulting from subsequent additions of a cross on position 1, a nought on 5, and again a cross on 7. Given the clear-cut connection between descriptions of this kind and the evidential situation that supports them, a formulation using lattice-connectives (which are extensional) is warranted.

Since every move in the game can also be seen as (in principle) precluding some other moves, we introduce extended descriptions of an evidential situation. This is obtained by enhancing the basic description with formulae obtained through the application of the following two rules:

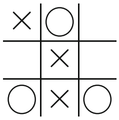
$$s_i \Vdash P_i a_j \ \& \ P_k \in (\mathsf{Position} \setminus \{P_i\}) \Rightarrow s_i \Vdash\!\sim P_k a_j \tag{1.10}$$

$$s_i \Vdash P_i a_j \ \& \ a_l \in (\mathsf{Turn} \setminus \{a_j\}) \Rightarrow s_i \Vdash\!\sim P_i a_l \tag{1.11}$$

As we seek to explain a modal space wherein $s \sqsubseteq s^*$ does not hold for arbitrary s, the set of all board-configurations cannot be restricted to those that respect the rules. Specifically, the set S includes several impossible configurations (most evidently boards containing overlapping noughts and crosses), while the subset $\{s : s \sqsubseteq s^*\}$ can be said to contain all and only those boards which result from possible games. By considering the rules given above as a means to encode some non-overlap rules within the descriptions themselves, impossible configurations are just those configurations which support an inconsistent description; impossibility and negation-inconsistency are made to coincide. The incorporation of such impossibilities is motivated by the paraconsistently inspired concern that within this example, a non-reglementary game should not be considered a trivial game (a game wherein each move is actually made), but only an impossible game (a game that no move can turn into a completed game). Alternatively, the set S can be thought of as the set of all boards someone who observes, but does not know the rules of the game can conceive of. Its subset $\{s : s \sqsubseteq s^*\}$, however, is still constrained by what can actually occur as a node in a game-tree.

So far, this only suffices to model evidential situations one can describe using lattice-connectives and negation only. A further extension for the group-connectives remains to be given. Consider the board-configuration and matching description in Fig. 1.1:

Fig. 1.1 Board-configuration



$s \Vdash P_1a_1 \sqcap P_7a_2 \sqcap P_5a_3 \sqcap P_9a_4 \sqcap P_8a_5 \sqcap P_2a_6$

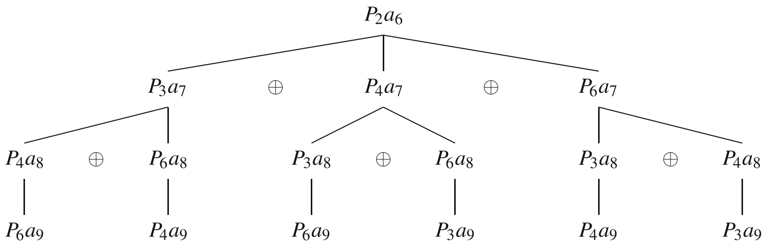


Fig. 1.2 Game-tree

Informally, we certainly would like to say that given this board, a sender who knows the game should be able to communicate more than what is true at s only (i.e. the closure of the description under $\sqcup$, $\sqcap$, and the additional rules (1.10) and (1.11)). For instance, it makes sense that $P_3a_7 \oplus P_4a_7 \oplus P_6a_7$, or even that $P_6a_7 \to (P_3a_8 \oplus P_4a_8)$. Hence, one could suggest to supplement the sender's evidential situation by means of a game-tree (see Fig. 1.2) depicting all constraints on the possible extensions of the actual board-configuration.

In the light of previous game-theoretical treatments of linear logic, it is an obvious move to explain the meaning of the group-disjunction $\oplus$ on the basis of moves a player can make in a game.[5] The novelty of the present approach is, therefore, situated on a different level; viz. the use of the *method of abstraction* [12] as a formal approach to the receiver's failure to discriminate between lattice and group-theoretical disjunctions—something related to what Humberstone calls logical discrimination [16].

Concretely, the modelling involves the introduction of two perspectives or levels of abstraction, depending only on the ability to discriminate between messages sent by a group of agents (where information available in the group need not be available in extensional form to any individual agent), and messages sent by a single agent (hence, available in extensional form). Additionally, we stipulate that the receiver's perspective is such that any message is perceived *as if* it originated from a single agent. The receiver, in other words, is modelled with respect to the lesser discriminating level of abstraction. We start, however, with the description of the first, more elaborated level.

Let, as said, the sender be a group of agents such that:

> [FACT] A first agent has access to an extended evidential situation and can only send messages insofar as they are supported by this evidence. Informally this corresponds to the actual game-board, and the application of (1.10) and (1.11). Evidently, every message supported by this board can be expressed using lattice connectives and negation only.

---

[5] Note, however, that the group-theoretical conjunction equally expresses a choice; namely, the choice the player himself has.

[CONS] A second group of agents has access to constraining information regarding the possible ways in which the actual board-configuration can be extended. This group of agents can only send messages through collaboration. For instance, if n such agents cover all possible moves $P_i a_{10-n}$ given the actual board, they are entitled to convey the message $P_i a_{10-n} \oplus \ldots \oplus P_j a_{10-n}$ where each disjunct is such a move.

Referring to the game-tree depicted above, n agents can send a message $P_i a_{10-n} \oplus \ldots \oplus P_j a_{10-n}$ iff any disjunct of the message labels an immediate successor of the root, and every label of an immediate successor of the root is a disjunct of the message. More generally, $n+1$ agents can send a message $P_i a_{9-n} \rightarrow (P_j a_{10-n} \oplus \ldots \oplus P_k a 10 - n)$ iff any disjunct labels an immediate successor of a node labelled with $P_i a_{9-b}$, and every label of an immediate successor of a node labelled with $P_i a_{9-b}$ is a disjunct.

First Example. $P_3 a_7 \sqcup \sim P_3 a_7$ and $P_3 a_7 \oplus \sim P_3 a_7$
Evaluating these messages, a disagreement between the notion of situated factual content and situated constraining content surfaces. Whereas the lattice-formulation can only be truly asserted on the condition that the seventh move has a definite value, the group-formulation is not bound by this requirement. Hence, the former does convey factual content, while the latter (if evaluated with respect to *Log* only) does not put any real constraint upon the possible evolutions of the game;[6] it is informationally empty.

Since we wish to understand genuine constraining content as extra-logical content, it is a normal and harmless assumption to evaluate constraining content with regard to logical constraints only. A better appreciation of this approach can be obtained in a second example.

Second Example. $P_6 a_7 \rightarrow (P_3 a_8 \oplus \sim P_3 a_8)$
Individuating the (constraining) content conveyed by this message, we use $CONT_c(P_3 a_8 \oplus \sim P_3 a_8 \mid P_6 a_7)$.

$$\begin{aligned} CONT_c(P_3 a_8 \oplus \sim P_3 a_8 \mid P_6 a_7) &= CONT_c(P_3 a_8 \rightarrow P_3 a_8 \mid P_6 a_7) \\ &= CONT_c(P_3 a_8 \mid P_6 a_7;\, P_3 a_8) \\ &= (P_3 a_8 \otimes P_6 a_7)^{\sqsubseteq}_{c_1;c_2} \setminus \{s : s \Vdash P_3 a_8 \otimes P_3 a_8 \otimes P_6 a_7\} \end{aligned}$$

Keeping in mind that $B \rightarrow (A \rightarrow A)$ should not be a theorem of relevant logic, it is to be expected that the constraining content of the message above is non-empty. Yet, since in the previous example we settled that a message of the form $P_i a_j \oplus \sim P_i a_j$ had to be informationally empty, the failure of that analysis for this new message needs to be made explicit. Specifically, the constraint against which the evaluation of the whole message is carried out needs to be brought to the surface. In general, theoremhood is judged at a logical situation or constraint. This is why, if we take *DTL* seriously, constraining content is to be judged against a logical constraint as

[6] Remark in that light that rules (1.10) and (1.11) effectively turn some rules of the game into logical rules.

well. It is, however, a mistake to infer from this insight that *any* appeal to a constraint is therefore an appeal to a logical constraint. For instance, the judgement that $B \rightarrow (A \rightarrow A)$ is not relevantly valid explicitly bears upon the possibility of evaluating the consequent $A \rightarrow A$ with respect to all constraints. This very feature, is recovered within the content-individuation described above by means of $(P_3a_8 \otimes P_6a_7)^{\sqsubseteq}_{c_1;c_2}$, a proposition that possibly exceeds *Log*. Comparing thus the following two individuations, the distinction can be explicitized.

$$CONT_c(A \mid A) = A^{\sqsubseteq}_c \setminus \{s : s \Vdash A \otimes A\} \tag{1.12}$$

$$CONT_c(A \oplus {\sim} A \mid B) = (A \otimes B)^{\sqsubseteq}_{c_1;c_2} \setminus \{s : s \Vdash A \otimes A \otimes B\} \tag{1.13}$$

Whereas (1.12) individuates the proportion of c-accessible A-situations that do not support A for the basic case where $c \in Log$, (1.13) individuates the proportion of $c_1; c_2$-accessible $A \oplus B$-situations that do not support A on the weaker assumption that $c_1 \in Log$. No matter how much one knows about c_1, the stronger assumption that $c_1; c_2 \in Log$ cannot be derived from the former, and consequently one cannot in general rule out the possibility of (1.13) being judged against an impossible constraint.[7] Even then, it seems that the weak assumption that $c_1 \in Log$ is all one needs to comply with *DTL*. While the former argument only settles the case on the formal level, the connection with the game interpretation is quickly made. Assume, for instance, that the receiver cannot rule out that the actual board-configuration, say s, is impossible. In such a case, even though it is natural for this receiver to judge the constraining content of a message with respect to a $c_1 \in Log$ it does not have the resources to rule out that the result of applying c_1 to s (i.e. any x such that $s \sqsubseteq_{c_1} x$), is possible too. But this is exactly what is required to settle that the composed channel $c_1; c_2$ is logical.

Trying, in turn, to formalise the less discriminating level of abstraction associated with the receiver's perspective, it is the distinction between $\sqcup$ and $\oplus$ which ought to disappear. The classical connectives, however, cannot serve that purpose since, (i) using the classical disjunction as a means to render a message for which $n > 1$ agents collaborated would turn constraining content into factual content, and (ii) using the classical disjunction to render a single-agent disjunction would erroneously assign some inferential strength to a purely factual message. To the receiver, the group of agents is just seen as a simulation of a single agent. Consequently, the incoming messages are, even though the result of a distributed system involving both single and collaborating agents, considered as a standard set of premises wherein the fine-grained structure derived from $(S, Log, \sqsubseteq_{\_}, *)$ is lost. Unfortunately, in such a case the content of a message cannot safely be individuated by simulating the receiver's perspective through a set of possible worlds only. There

[7] Let $c = c_1; c_2$ where $c_1 \in Log$. Assume for reductio that $c \in Log$ too. By $c_1 \sqsubseteq_{c_2} c$, we have that $c_2 \sqsubseteq_{c_1} c$, and since $c_1 \in Log$ we also have that $c_2 \sqsubseteq c$. But then, given our assumption that $c \in Log$, it must at least hold that $c_2 \sqsubseteq c_2^*$. But since c_2 can be any element of S, the latter should not hold in general.

simply is no warrant for assuming the existence of a non-empty set of possible worlds or a *worldly perspective*. In other words, the previously made remark that a worldly perspective need not exist remains, even after the introduction of a notion of constraining content.

The level of abstraction that duly incorporates the ability of discriminating messages sent by one or more agents is best rendered using both the lattice and group-connectives. Informational content is, with respect to that level of abstraction, correctly rendered using two unambiguous notions of content. Namely, $CONT_c$ for the cases described under [CONS] and $CONT_s$ for the cases described under [FACT]. Coming to the less fine-grained receiver's perspective, the logical discrimination assumed at the previous level is totally lost. Yet, the classical collapse of previously distinct connectives may not apply as a rendering of this reduced discrimination. All one can say is that unbeknownst to the receiver, content is effectively governed by $CONT_s$ and $CONT_c$. These considerations do show that disjunctive messages are sometimes rightly described by means of the ambiguous connective '`or`', while their content should not necessarily be governed by their classical properties. In that sense, the receiver's perspective we have modelled, implicitly recaptures what Paoli calls the ambiguity of our natural language *or* [23].

As such, the framework outlined in this section suffices to tackle three subsequent objections to substructural logic, logical pluralism, and more generally the individuations of informational content based on them.

1.5 Three Objections Revisited

1.5.1 Burgess' Objection

In his 'Relevance: A Fallacy' [10], Burgess fiercely argues against the distinction between $\sqcup$ and $\oplus$ proposed by some relevantists, and especially against their contention that each apparently valid instance of $DS^{\texttt{or}}$ could be recast as an application of $DS^{\oplus}$. Upholding that, if not simply ad hoc, the distinction between lattice and group-theoretical connectives could be shown to reduce to a mere subjectivization of relevance which rests on a confusion of implication with inference, Burgess is not prepared to accept this distinction as a logical distinction.

> Relevantism would reduce to the position that (IA) [$DS^{\texttt{or}}$] is valid when and only when one's grounds for asserting $p \vee q$ are something other than the simple knowledge that q. Such a position, however, looks suspiciously like a confusion of the criteria for the *validity* of a form of argument with the criteria for its *utility*, a confusion of logic with epistemology. (...)
>
> Thus if [Anderson and Belnap] intend by 'relevance' something less than objective, they are highly remiss in failing to alert their readers to the fact; while if 'relevance' is supposed to be impersonal, then the claim that the relevantistic position is (even in a weak sense) compatible with commonsense and accepted mathematical practice succumbs to the counterexamples presented above. [10, p. 103]

Accepting Burgess' claims, one can only conclude that there is no viable notion of informational content beyond that of worldly content. The resulting position is that of a classical monist, a view that squarely contradicts the findings of the preceding sections.

What allows us to escape this conclusion, is the adoption of a broadly Dretskian attitude when reformulating this first objection in terms of content and content-containment. It is therefore hardly surprising that our response departs in many ways from earlier defences of relevantly inspired recaptures of the disjunctive syllogism [20, 21, 28–30].

The crucial point of our argument is already implicit in the previously made assertion that the content of a message 'A `or` B' is, even in those cases where the receiver lacks the resources to correctly disambiguate it, either correctly individuated using $CONT_s$ or using $CONT_c$. In short: the content a message conveys is determinate, and this is so independently of the receiver's actual knowledge of that content. Such a remark is reminiscent of many points in Dretske's 'Knowledge and The Flow of Information'.

> The explanation of this 'paradox' lies in the fact that the information (...) can be communicated over a channel without the receiver's knowing (or believing) that the channel is in a state such as to transmit this information. The receiver may be quite ignorant of the particular mechanisms responsible for the delivery of information—holding no beliefs, one way or the other, about the conditions on which the signal depends.
>
> Information (and therefore knowledge) about a source depends on a reliable system of communication between a source and receiver—not on whether it is known to be reliable. [11, chap. V]

Reconsidering Burgess' objection in this light (and not only with respect to the view he explicitly challenges, viz. Anderson and Belnap's), one can easily point to a first confusion on his side. To wit, if the distinction between $A \sqcup B$ and $A \oplus B$ were just a matter of subjectivity—of knowledge or the lack of knowledge of either A or B—then the ability to discriminate between them had better be assigned to the receiver. Yet, this is exactly what our model denies. The information explicitly available to the receiver is mediated by a level of abstraction whose main characteristic is its failure to discriminate between messages sent by one or by more agents. The distinction between lattice and group-theoretical connectives is therefore equally unavailable to the receiver. Whether the information conveyed by the ambiguous message 'A `or` B' is factual or constraining is judged independently from the receiver's knowledge of the truth of A or B. Thus, if the model presented in the previous section adequately reflects the distinction between lattice and group-theoretical connectives, this very distinction cannot be explained in terms of the subjectivity of the receiver. Yet, since Burgess mentions the grounds for asserting a disjunction instead of the resources to recognise it, we need more to meet his challenge.

More important therefore, is the fact that the distinction cannot be recast in terms of subjectivity of the sender either. Given an accurate use of the distinct levels of abstraction at work, it can be shown that a reduction to the information available to the sender is, if not straightforwardly based on a confusion, then at least based on an incomplete understanding of the context of communication. The tempting mistake rests on the intuitive truth of the following claim: a truthful message 'A `or` B' conveys the information associated with $A \sqcup B$ iff the sender could—given the evidence presently at its avail—have sent a more informative message that entails either A or B; if not, then it conveys the content associated with $A \oplus B$. Quite rightly, the ability of conveying the information associated with one or another disambiguation can be reduced to the ability to convey the information associated with its subformulae. Nevertheless it only provides an accurate description given the assumption that messages are perceived as coming from a single agent (as we have seen, the receiver's perspective). This assumption, however, amounts to reducing our understanding of the context of communication to a single level of abstraction.

The view we wish to defend encompasses two claims. First, the content effectively conveyed by a message (alternatively, what counts as its correct disambiguation) is a determinate matter in that it is independent from the receiver's previously acquired information. This defuses Burgess' objection that the relevantist confuses inference with implication. Secondly, what content is effectively conveyed can only be recast in simple epistemological terms if the sender is perceived as a single agent, and—as discussed above—the latter depends on the adoption of the lesser discriminating level of abstraction. As soon as the 'sender' is recognised as a group of agents that possibly need to collaborate to send a message,[8] the epistemological explanation for the distinction is provided with a second, not purely epistemic ground. This, in the end, shows the non adhocness of the distinction.

A final objection to Burgess' criticism refers to his failure to recognise the possibility that information might not only be distributed over distinct agents, but might be so unbeknownst to the receiver of the message that conveys this information. Since, as argued in a previous section, the failure to discriminate the fine-grained structure of the information cannot safely be described using a worldly perspective only (i.e. classical logic), a reason for not being a classical monist is obtained.

1.5.2 Read's Objection

Being one of the most consistent advocates of the distinction between lattice and group-theoretical connectives, any objection to informational pluralism derived from Read's position is bound to be diametrically opposed to the one discussed above. Despite the vagueness of the label, we shall refer to his position as a *relevantist's monism*, and devote our attention to the consequences of his rejection of both classical logic and a classical meta-theory. As voiced on many occasions, it is Read's opinion that the Routley-Meyer semantics for relevant logic is—as Meyer himself claimed before—indeed a *gentile semantics* [30]. It uses an extensional language to reconstruct an essentially intensional consequence relation. More recently,

---

[8] Remember that our appeal to multiple agents is itself an artefact we use to account for the real phenomenon under consideration: the distributed nature of information (on that topic, see also [7]).

in objecting to logical pluralism, the same diagnosis was advanced as a central part of his argument: 'If Beall and Restall insist on doing semantics classically, then they are just classical logicians' who think of non-classical logics as incomplete instead of truly rivalling logics [31, pp. 207–09]. Their pluralism depends more on a failure to really reject classical logic, than on a robust position in itself.

On the face of it, the informational pluralism outlined above falls prey to both sides of Read's objection. Not only does the model presented in Section 1.4 take the Routley-Meyer semantics very seriously, but its explicit endorsement of a notion of *worldly content* boils down to an overt acceptance of classical logic. Our defence against a relevantist's monism shall comprise two parts; a first part to defend our modelling, and a second one specifically in favour of the indispensability of a notion of *worldly content*.

Regarding the first, a fairly recent tradition of formulating intuitive interpretations for formal semantics previously deemed opaque, confers some initial plausibility to our enterprise. The fact that two such models [17, 32] fall back on situation-semantics and information-flow furthermore suggests the viability of an information-based reading of the semantics. Yet, it speaks in our model's favour that by avoiding the explicit references to information-flow (according to some an equally obscure notion) both [17] and [32] make use of, it cannot as easily be dismissed either. After all, our initial aim was not as much to provide an intuitive interpretation of the semantics of relevant logic as well as an attempt to devise a more fine-grained account of content and content-containment.[9] In that perspective, the model based on communicating agents comes first, and—however it makes the distinction between lattice and group-theoretical connectives less obscure—its main purpose is not to defend an extensional semantics as a non-gentile semantics. Even though one might conclude that it makes the Routley-Meyer semantics more respectable, this conclusion is largely immaterial to the pluralist position about informational content defended here. The pluralism we take to be unavoidable follows from the interpretation in terms of communicating agents, not from the fact that an extensional semantics is used.

Since it touches upon the core of our pluralism, Read's rejection of classical logic [30] is more troublesome. As he has it, relevant logic discriminates between two accounts of consequence, one material and one relevant. Classical logic, to the contrary, conflates them. So far, we cannot but agree: the first four sections of this paper recognise that from a worldly perspective constraining and factual situated content fully coincide. Yet, beyond this point the disagreement with Read's monism is blatant. To see how this affects our position, we need to go back to the intuition behind *DTL*: the fact that informational content and logical consequence should be considered as dual. If Read is right in rejecting both material and classical consequence, of the three measures we proposed only $CONT_c$ survives. As from a monist perspective they cannot comply with *DTL*, the material $CONT_s$ and the classically

[9] As suggested by Greg Restall (pc), the model of Section 1.4 might in a sense be closer to the approach in his 'Modelling Truthmaking' [33].

minded $CONT_w$ only count as formally correct precisifications of *GIRP*—mere artefacts of our usage of a *case*-based semantics. This leaves us with a number of distinct issues to resolve. Not only do we need to establish the usefulness of $CONT_s$ and $CONT_w$ (*qua* kinds of content), but a coherent case in favour of *DTL* remains to be presented for each of them too.

As an argument for the relevance of $CONT_s$, it should be sufficient to recall the previously suggested interpretation of situated factual content. Viz. A is non-empty only if A could not have been communicated by a sender in any situation. Recognising the partiality of information-states, and the distributed nature of information in a complex environment, the notion of situated factual content can be considered a useful tool to cope with these features.

Yet, as we reformulate Read's rejection of material consequence in informational terms, we do not object directly to the former consideration, as well as to its compliance with the *DTL*-intuition. Extending our language with the lattice-theoretical constant t (which should not be confused with the truth-value),[10] this concern can be given a more precise formulation. Defining the conditional content using the lattice conjunction, and the constant t with $s \Vdash t$ for all $s \in S$, we can exploit the material equivalence of A and $t \rightsquigarrow A$ (where $\rightsquigarrow$ is a lattice implication) as follows.

$$\begin{aligned} CONT_s(A \mid t) &= CONT_s(A \sqcap t) \setminus CONT_s(t) \\ &= (S \setminus \{s : s \Vdash A \sqcap t\}) \setminus (S \setminus \{s : s \Vdash t\}) \\ &= S \setminus \{s : s \Vdash A\} \\ &= CONT_s(A) \end{aligned}$$

This, in turn, contributes to a suitable perspective that gives situated factual content an interpretation in terms of material consequence: it connects situated factual content to derivability in an arbitrary situation. Unfortunately, derivability at an arbitrary situation is hardly better than truth. It does not give us a better grip on the logicality of $CONT_s$. As impossible situations cannot be excluded, nothing is true in all situations; derivability fails, exactly like truth does, to behave in a logically constrained way.

In the end, we do not quite see how this conclusion can be avoided. Even more, by separating the factual from the constraining content of a message we may have broken the previously existing connection between content-containment or consequence and content simpliciter. It is hard to see how we can at the same time draw a line between facts and constraints, and provide more than a fairly thin case for the logicality of the former. Such a thin case can take two forms. The first one, is to accept $CONT_s(A \mid B)$ as a degenerate, but still logical account of content-containment. This line is, given its dependence on the lattice-constant t for connecting content and content-containment, not our preferred option (see also [25]).

[10] An often-used symbol for the lattice-theoretical constant is $\top$ (top). I prefer to use t and keep $\top$ for the classical truth-constant.

The second case makes an appeal to the behaviour of $CONT_s$ with respect to, on the one hand, possible situations, and, on the other hand, logical situations. As mentioned, a situation is impossible just when it supports a contradiction. Conversely, a situation s is possible if it fails to support any contradiction; i.e. when it holds that $s \sqsubseteq s^*$. This, one could suggest, makes possible situations apt for logical behaviour: they never explicitly deny a classical theorem. However, as hinted upon in Section 1.2, the set of possible situations is not upwardly closed under $\sqsubseteq$, and the property of not denying a classical theorem can therefore not be expressed as a proposition. This is exactly where [21] fails in restoring the material disjunctive syllogism using consistency as a premise [18, pp. 183–84]. In that respect, logical situations are more robust: the set *Log* is upwardly closed, and hence 'being logical' expresses a proposition. Thus, if 1 is the group-theoretical constant satisfied by all $s \in Log$, we know that if $s \Vdash 1$, we also have $s' \Vdash 1$ for all $s \sqsubseteq s'$, and, as a consequence, we also know that if $s \in Log$, then $s = s^*$; persistently failing to deny a classical theorem just means asserting that theorem.[11]

On how this affects the notion of factual content, we can make the following comments. If s is excluded by A ($s \in CONT_s(A)$), and A is a classical theorem, then s is only persistently excluded if s is impossible. Generally, the possible situations individuated by $CONT_s(A)$ are just those situations that have not yet decided whether A, and might thus be extended either way: they only *weakly exclude* A. Put differently, $CONT_s(A)$ individuates those situations wherein one cannot soundly infer A on the basis of explicitly availably information only. For an arbitrary situation, it means that nothing at all can thus be inferred.

Alternatively intersecting $CONT_s(A)$ with the set of situations satisfying 1 (i.e. the logical situations), we can exploit the weak equivalence of A and $1 \rightsquigarrow A$ and show the duality of situated factual content and enthymematic derivability;[12] that is, we assume 1 to be a suppressed but obviously true antecedent [4, §35]; [22, pp. 78–79].

$$\begin{aligned}
CONT_s(A \mid 1) &= CONT_s(A \sqcap 1) \setminus CONT_s(1) \\
&= (S \setminus \{s : s \Vdash A \sqcap 1\}) \setminus (S \setminus \{s : s \Vdash 1\}) \\
&= \{s : s \Vdash 1\} \setminus \{s : s \Vdash A \sqcap 1\} \\
&= \{s : s = s^*\} \setminus \{s : s = s^* \ \& \ s \Vdash A\} \\
&= CONT_w(A)
\end{aligned}$$

Since given the explicit assumption that 1, the situated factual content of A is just the worldly content of A, the robust logicality of $CONT_s$ cannot be regarded independently from $CONT_w$. At first sight, this is just additional evidence for Read's point

[11] If $s \Vdash 1$ then s is consistent too, that is $s \sqsubseteq s^*$. Consequently, $s^* \Vdash 1$ holds too, and by the same token $s^* \sqsubseteq s^{**}$. Since $s = s^{**}$, $s^* \sqsubseteq s$, and hence $s = s^*$.

[12] Traditionally, an enthymeme is an argument with an unstated or suppressed assumption. In the relevantist tradition it is common to recapture classical reasoning enthymematically by treating consistency (in some or other form) as an unstated assumption.

of view [31]: the classical account is predominant, and situated factual content is just a crippled classical account. By providing an independent reason of accepting the latter, the slide to worldly content is shown to be harmless for the independence of the situated account.

To motivate the respectability of a notion of worldly content, we need to sketch a situation where it provides the most accurate account of informational content. To do so, we rely on a strategy outlined in [2] and provide a plausible interpretation of a worldly perspective. Let us go back to our initial example, and assume that right after the second move in the game was made (see Fig. 1.2), the sender wrote the message 'P_3a_7 or $\sim P_3a_7$' on a slip of paper. Imagine that this message remained obscured to an hypothetical receiver until right after the game was duly completed. If so, one cannot but conclude that (no matter how the game was completed) the message is devoid of any content. However, since we could have come to the same conclusion using $CONT_c$, this is hardly sufficient to settle the correctness of $CONT_w$. The dispensability of $CONT_w$ is exactly what we could have expected on the basis of what many proponents of relevant logic have claimed: relevant logic is self-sufficient, it validates all the classical inferences while avoiding its confusions. What we do miss in using $CONT_c$ as an all-purpose measure, is the fact that from a worldly perspective the content conveyed by this message is as factual as it can be. More generally, in our use of $CONT_w$ we stress that, irrespectively of whether from a situated perspective a message conveys factual or constraining content, from a worldly perspective it unavoidably conveys both. What our notion of worldly content accounts for, is the additional consideration that, from a worldly perspective only, factual and constraining content actually do coincide. Specifically, their collapse is not a matter of equivocation, but of a lesser—yet perfectly adequate—discriminatory power: it does not rest on a confusion of factual and constraining content, but on their actual inseparability.

Here too, one can object that this collapse rests upon weak equivalences of lattice and group-disambiguations only. For both kinds of content to be equivalent, an additional truth must be assumed. But then, if the functioning of worldly content—including the collapse of facts and constraints—can be simulated by means of an additional premise, why include an independent account of worldly content? This question is answered in the next section as a reply to a monist objection due to Priest.

1.5.3 Priest's Objection

As anyone supporting the correctness of a sub-classical logic, both monists and pluralists face the problem of explaining the apparent validity of inferences their preferred logic deems invalid. Obviously, the monist cannot appeal to stronger but equally good logics to cope with this issue, and must therefore look for other solutions. It is very likely that, as Read [30] shows, an appeal to unambiguous connectives can do most, if not all the work a monist requires of it. Yet, if our argument in favour of a less-discriminating (worldly) perspective has any ground, the monist must come up with additional methods to accommodate this last phenomenon.

As suggested above, a standard monist strategy can consist in adding new, domain-specific premises.

> Now, just as an intuitionist may use what amounts to classical logic when reasoning about finite situations, so a paraconsistent logician may use what amounts to classical logic given appropriate information about the domain. For example, sufficient information is that for all α, $(\alpha \wedge \neg\alpha) \rightarrow \bot$, (…). [27, p. 28]

Using a system weaker than Priest's preferred logic, the additional premise we must appeal to is the group-theoretical constant 1. Despite the apparent effectiveness of this approach, we do not think it can ultimately replace an appeal to worldly content. To that effect, we advance two arguments, a proof-theoretical and a semantic one.

The proof-theoretical argument rests upon the fact that to prove the equivalence of lattice and group-theoretical connectives, the inclusion of 1 does not help. Adding it as a premise does warrant that *weak derivability* is just fine, but what it cannot do, is make the structural rules required for a weak derivability relation obsolete.

Semantically, however, we have already seen that adding 1 as an antecedent suffices to get $CONT_s(A \mid 1) = CONT_w(A)$. To bridge the gap between constraining and worldly content, a restriction to logical constraints as well as the assumption that A does not contain any group-theoretical connectives is additionally required.

$$\begin{aligned} CONT_c(A \mid 1) &= 1_c^{\sqsubseteq} \setminus \{s : s \Vdash 1 \otimes A\} \\ &= \{s_i : \exists s(s \Vdash 1 \;\&\; \forall A(c \Vdash A \Rightarrow s_i \Vdash 1 \otimes A))\} \setminus \{s : s \Vdash 1 \otimes A\} \\ (\text{if } c \in Log) &= \{s : s = s^*\} \setminus \{s : s = s^* \;\&\; s \Vdash A\} \\ &= CONT_w(A) \end{aligned}$$

Still, the specific semantic argument we want to present, does not directly depend on these formal considerations, but remains closer to the interpretation we gave in terms of communicating agents. It crucially hangs on the specific way we flesh out the distinction between situated and worldly content.

For 1 to be considered an additional premise within a context of communication, it can only be introduced as an explicitly sent message. This, however, requires the sender to have sufficient information at its disposal to assert 1, and the receiver to have the ability to recognise it as such. As should be clear by now, neither of these conditions can actually be fulfilled. All the sender and the receiver can be said to hold, is the weaker *non-persistent* consistency-premise. Recasting Gillies' strategy of dealing with epistemic modals [13, 14], consistency is better considered the result of a successful test upon (or a global property of) one's actual state of information rather than a genuine piece of information itself. Quite like Mortensen's consistency premise [21], the result of such a test cannot be preserved in one's later states of information. As these specific limitations show that there is no reliable test for the satisfaction of 1, no message can reliably convey the information required to settle its truth either.

So far, this only establishes that for a situated agent no message can persistently convey the information that there exists a worldly perspective upon the context of communication it is part of. It nevertheless remains possible that if a worldly perspective exists, the correct way of assessing content and content-containment from that perspective crucially depends upon the availability of an additional premise. Upon closer inspection, we think this option should equally be dismissed. Surely, adding a premise apparently gets you right there, but it also fails to acknowledge an important difference between the situated and the worldly perspective. What we specifically need to point at, is the specific role played by the additional premise. That is, it explicitly states the information a situated agent requires to assess the content of a message *as if* it were evaluated from a worldly perspective. This strategy enables a situated receiver to simulate the worldly perspective by assuming 1 to be true, not by knowing it to be true. Put differently, there is a gap between correctly assuming that all the messages one receives originate from a single agent, and the fact that there is only one agent who sends messages.

The two main properties of this monist strategy, are now clear: the inclusion of the new premise proceeds explicitly, and by assumption only. But if this is all the monist can advance, the following dilemma is hardly avoidable. Either the monist cannot really account for the worldly perspective, or otherwise informational pluralism cannot be avoided by the logical monist. The first half is fairly trivial: that there is a worldly perspective just means that 1 is true, not that it is merely assumed to be true. But then, if 1 can be true while no one can explicitly be informed of its truth, it can only be concluded that $CONT_w$ correctly assesses the worldly content of a message because it implicitly incorporates the truth of 1. Yet, this is exactly what the pluralist with respect to informational content claims.

1.6 Conclusion: A Realist's Pluralism

To conclude, we still have to explain what the *realism* in the commitment to pluralism stands for. Basically, if logic deals with inferences based on explicitly available information only, a monist can show that he accommodates for all extra-logical reasoning (most likely classical inference-rules that go beyond what can be achieved through disambiguation only) by the explicit addition of supplementary premises. Yet, if we speak of content and content-containment along realist lines, we often have to make an appeal to a relation of content-containment or nesting that the receiver of a message does not know of. In such cases, the relevant work cannot any longer be done by explicitly adding premises; only a logical solution is acceptable.

Specifically, this shows that the usual monist strategy of pretending that one only needs to pay attention to the premises no longer works if our focus is content, and content-containment. This insight, forms the basis of a realist commitment to logical pluralism.[13]

[13] This paper was originally written in the Summer of 2006, and complements [2]. Both these papers argue for logical and informational pluralism in the same purely model-theoreric fashion; a method I would no longer rely on in the same way as I did. Yet, I've chosen not to actualise the present paper to reflect these changes, but to stick to the original version. For the same reason, references to more recent literature on the topic of logical pluralism haven't been included either.

References

1. Allo, P. 2007a. "Formalising Semantic Information. Lessons from Logical Pluralism." In *Computing, Philosophy, and Cognitive Science*, edited by G. Dodig Crnkovic and S. Stuart, 41–52. Cambridge, MA: Cambridge Scholars Press.
2. Allo, P. 2007b. "Logical Pluralism and Semantic Information." *Journal of Philosophical Logic* 36:659–94.
3. Anderson, A. R., and N. D. Belnap. 1975. *Entailment. The Logic of Relevance and Necessity (Vol. I)*. Princeton, NJ: Princeton University Press.
4. Anderson, A. R., N. D. Belnap, and J. M. Dunn. 1992. *Entailment. The Logic of Relevance and Necessity (Vol II).* Princeton, NJ: Princeton University Press.
5. Bar-Hillel, Y., and R. Carnap. 1952. *An Outline of a Theory of Semantic Information*. Cambridge, MA: MIT, Technical Report 247. Reprinted in: Bar-Hillel, Y. 1964. *Language and Information. Selected Essays on Their Theory and Application.* London: Addison-Wesley.
6. Barwise, J. 1997. "Information and Impossibilities." *Notre Dame Journal of Formal Logic* 38:488–515.
7. Barwise, J., and J. Seligman. 1997. *Information Flow: The Logic of Distributed Systems*. Cambridge, MA: Cambridge University Press.
8. Beall, J. C., and G. Restall. 2000. "Logical Pluralism." *Australasian Journal of Philosophy* 78:475–93.
9. Beall, J. C., and G. Restall. 2006. *Logical Pluralism*. Oxford: Oxford University Press.
10. Burgess, J. P. 1981. "Relevance: A Fallacy." *Notre Dame Journal of Formal Logic* 22:97–104.
11. Dretske, F. 1999. *Knowledge and the Flow of Information*. Stanford, CA: CSLI.
12. Floridi, L. 2008. "The Method of Levels of Abstraction." *Minds and Machines* 18:303–29.
13. Gillies, A. S. 2004. "Epistemic Conditionals and Conditional Epistemics." *Noûs* 38:585–616.
14. Gillies, A. S. 2006. "What Might Be the Case After a Change in View." *Journal of Philosophical Logic* 35:117–45.
15. Hanson, W. H. 1980. "First-Degree Entailments and Information." *Notre Dame Journal of Formal Logic* 21:659–71.
16. Humberstone, I. L. 2005. "Logical Discrimination." In *Logica Universalis*, edited by J.-Y. Béziau, 207–28. Basel: Birkhäuser.
17. Mares, E. 1997. "Relevant Logic and the Theory of Information." *Synthese* 109:345–60.
18. Mares, E. 2004. *Relevant Logic—A Philosophical Interpretation*. Cambridge, MA: Cambridge University Press.
19. Mares, E. 2006. "Relevant Logic, Probabilistic Information, and Conditionals." *Logique and Analyse* 49:399–411.
20. Mortensen, C. 1983. "The Validity of Disjunctive Syllogism Is Not So Easily Proved." *Notre Dame Journal of Formal Logic* 24:35–40.
21. Mortensen, C. 1986. "Reply to Burgess and Read." *Notre Dame Journal of Formal Logic* 27:196–200.
22. Paoli, F. 2002. *Substructural Logics a Primer*. Dordrecht: Kluwer.
23. Paoli, F. 2003. "Quine and Slater on Paraconsistency and Deviance." *Journal of Philosophical Logic* 32:531–48.
24. Paoli, F. 2005. "The Ambiguity of Quantifiers." *Philosophical Studies* 124:313–30.
25. Paoli, F. 2007. "Implicational Paradoxes and the Meaning of Logical Constants." *Australasian Journal of Philosophy* 85:553–79.
26. Priest, G., ed. 1997. *Notre Dame Journal of Formal Logic* 38(4). (Special Issue on Impossible Worlds).
27. Priest, G. 2001b. "Logic: One or Many?" In *Logical Consequence: Rival Approaches*, edited by J. Woods and B. Brown, 23–38. Stanmore: Hermes.
28. Read, S. 1981. "What Is Wrong with Disjunctive Syllogism?" *Analysis* 41:66–70.
29. Read, S. 1983. "Burgess on Relevance: A Fallacy Indeed." *Notre Dame Journal of Formal Logic* 24:473–81.

30. Read, S. 1988. *Relevant Logic. A Philosophical Examination of Inference*. Oxford: Basil Blackwell.
31. Read, S. 2006. "Monism: The One True Logic." In *A Logical Approach to Philosophy*, edited by D. DeVidi and T. Keynon, 193–209. Dordrecht: Springer.
32. Restall, G. 1995. "Information Flow and Relevant Logic." In *Logic, Language and Computation: The 1994 Moraga Proceedings*, edited by J. Seligman and D. Westerståhl, 463–77. Stanford: CSLI Press.
33. Restall, G. 2000. "Modelling Truthmaking." *Logique and Analyse* 43:211–30.
34. Sequoiah-Grayson, S. 2006. "Information Flow and Impossible Situations." *Logique and Analyse* 49:371–98.

Chapter 2
My Own Truth

Pathologies of Self-Reference and Relative Truth

Alexandre Billon

2.1 Introduction

I'm a foresighted person. Long time ago I wrote the following sentence in my notebook

- ($\star$): Tr($\star$)

  (this should read: ($\star$) is true, and ($\star$) names this very sentence)

Some philosophers think that this sentences is false [15]. Others believe that it is true [21]. Yet others suggest that it is neither true nor false [24].[1] I believe those philosophers are all wrong, although they might have been right for a while. In fact, I believe that sentences of this type can say many different kinds of things. I even believe that one can change what such a sentence says *on a given occasion of use* very easily, and almost at will. For example, right now I am concentrating, and I say that, among other things, ($\star$) states a proof of Fermat's last theorem. Thus, I can rightly say that I wrote such a proof on my notebook: a sentence, ($\star$), of which it is now the case that it states such a proof, even if it has not always been the case. Still, I produced the proof. Writing ($\star$) on my notebook was a wise thing to do.

Truth-Tellers

Let me explain. I will call 'truth-tellers' tokens of the Truth-Teller, that is, token sentences which say of themselves that they are true, and nothing but that. We will assume that the language is fixed, so that when we say that a sentence is 'true', we will mean 'true in English'.

---

[1] Goldstein [7] does not explicitly consider sentences of this kind (truth-tellers) but he considers sibling problematic sentences (such as the 'Open Pair', cf. below) and claims that they do not say anything. Read [18] claims that they do not say anything but he later disavowed that claim [17].

A. Billon (✉)
Université Lille Nord de France, F-59000 Lille, France

UdL3, STL, CNRS, UMR 8163, F-59653 Villeneuve d'Ascq, France
e-mail: alexandre.billon@m4x.org

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_2, © Springer Science+Business Media B.V. 2012

Context-Sensitivity

By ‘sentence,’ I mean a mere grapheme or phoneme of the English language. We should distinguish sentences from what they state (their content, what they say, the proposition they express), for even a single token sentence can state many different things when used on different occasions. For example, while visiting a sleeping zoo at night, I can write ‘this one is dangerous’ on a piece of paper:

- (Z): ‘this one is dangerous’

(This should read: (Z) names the token of the type sentence ‘this one is dangerous’ above. Without the single quotes the sentence would have been used rather than just mentioned, so it should have read: this one is dangerous, and (Z) names the token the type sentence above)

I can show (Z) to my friend to tell him how dangerous the sleeping animal is that we are looking at. On this occasion of use O, (Z) states exactly that (say) the lion at such and such location is dangerous:

- $(Z)_O \mid$ the lion at such and such location is dangerous.

(‘$(Z)_O$’ stands for the sentence use of (Z) on occasion O, ‘$|$’ stands for ‘states exactly’. I will omit the mention of the occasion when it is irrelevant).

If a little later, say on occasion O’, I point at another animal and show the same piece of paper again, I will be making a different statement with the very same token sentence.[2]

- $(Z)_{O'} \mid$ the tiger at such and such location is dangerous.

Accordingly, the two different uses of the sentence could have different truth values. We will say that a token or type sentence is context-sensitive if it does not have a constant truth value across contexts.

Indexicalist Accounts of Context-Sensitivity

In cases like the one above, the context-sensitivity can be accounted for in indexical terms. In particular, what a sentence type or token says varies functionally with occasions of use. In other words, (i) what it says varies with occasions of use; (ii) but what it says *on a given occasion of use* is fixed and determined, (iii) so that taken together, a sentence of our language and an occasion of use of that sentence determine (at most) a truth value (at a world). Yet, one should be wary not to equate context-sensitivity with indexicality.

[2] Occasions of use must solve all indexical ambiguities so they must determine a world, a location, a time, a subject, etc. Instead of ‘occasions of use’ I could have used Kaplan’s term ‘context of utterance’ or Lewis’ ‘context’.

Relativist Accounts of Context-Sensitivity

There may be other sources of context-sensitivity. Some have argued that although they are context-sensitive in that their truth value varies with the occasions of use, temporal or first-person sentences have the very same content on different occasions of use. Such sentences would express so-called temporal or perspectival statements: statements that depend on the context of their use. Following MacFarlane [13], we can call this view nonindexical contextualism.[3] Both indexical and nonindexical contextualism claim that sentences are use-sensitive. They both claim that the truth value of a sentence varies functionally with the *context of its use* (indexical contextualism further argues that the content of the sentence varies in the same way). Both should be sharply distinguished from the thesis to the effect that the truth value of given sentence is not fixed and determined once and for all, but depends on the *context of its assessment*. It has been argued, for example, that the truth value of sentences dealing with future contingents depends on the time at which they are assessed [11], that sentences dealing with matters of taste or ascribing knowledge or justification depend on the standards adopted to assess it, etc. (see [9] for a synthesis). We will call relativist the claim that the truth value of given sentence use varies with the context in which it is assessed. For a given sentence use, such a context of assessment should determine all the parameters that are relevant to determine the truth value of that sentence. So for example, a relativist about sentences dealing with future contingents will typically represent contexts of assessment by times: a relativist about sentences dealing with matters of taste will typically represent contexts of assessment as standards of taste, etc. Like contextualism, relativism comes in two guises: *expressive relativists* claim that the sentence says different things in different contexts of assessment, and *statement-relativists* (in MacFarlane's [12] terminology, propositional relativists) claim that the sentence expresses the same statement in different contexts of assessment but that this statement is relative to its context of assessment: it has a different truth value in different contexts of assessment (Fig. 2.1).

	use sensitivity	**assessment sensitivity**
variable content	indexical contextualism	expressive relativism
constant content	nonindexical contextualism	statement relativism

Fig. 2.1 Various accounts of context-sensitivity

[3] Because temporal and perspectival statements are sometimes said to be 'relative' (to the time of their use, to the subject who uses them) nonindexical relativism is sometimes called relativism (Recanati [19] speaks of moderate relativism, Kölbel [9] of relativism). We will, however, reserve the word 'relativism' for dependence of the truth value on the context of *assessment*.

Overview

Returning now to the truth-teller we can make our claims more precise. The following is usually assumed without further ado

- **Uniformity**: All truth-tellers have the same semantic value (or the same lack thereof). They are not context-sensitive.
- **Absoluteness**: A fortiori, truth-tellers are not relative. The truth value of a given truth-teller use is not relative to the context in which it is assessed.

In what follows, I will show that these two claims are wrong. The Truth-Teller exhibits a radical form of *relativity*. Against *uniformity*, different tokens of the Truth-Teller must have different truth values and imply different things. Indeed, we will see that there is virtually no limit to what truth-tellers can imply (Section 2.2). Against *absoluteness*, I will then give a stronger argument to the effect that one can change the truth value of a given token truth-teller *on a given occasion of use* very easily, by changing the context of its assessment (Section 2.3). Finally I will show how this resolution of the Truth-Teller in terms of relative truth can be extended to other semantic pathologies of self-reference such as the Liar and the Open Pair (Section 2.4), and draw a broader moral about self-reference and intentionality (Section 2.5).

2.2 The Truth-Teller Is Context-Sensitive

In order to show that various token truth-tellers must have different truth values I will construct two sentences and I will show that although (i) they are both truth-tellers, (ii) they must nevertheless differ in truth value. Consider for example (1) and (2), where ‘p’ and ‘r’ are arbitrary sentences:

- (1): ‘p and not False((1))’
- (2): ‘r or not False ((2))’

It should first be noted that if we choose ‘p’ so that it is bivalent—that is, either true or false, not gappy—then (1) will be bivalent as well. If (1) were indeed gappy, it would not be false. Accordingly, ‘not False((1))’ would be true. But then (1), that is ‘p and not False((1))’, would have the same truth value as ‘p’. As *ex hypothesis* ‘p’ is not gappy, (1) would not be gappy either, which contradicts our hypothesis. So (1) cannot be gappy if ‘p’ is not.[4] From now on we will consider that ‘p’, and accordingly (1), are both bivalent.

Now it takes little work to see that because of its self-referential character, (1) is indeed a truth-teller: it says of itself that it is true, and nothing but that. Notice first that, on any occasion O, $\mathrm{Tr}((1)_O) \Leftrightarrow$ (p and not $\mathrm{False}((1))_O$):

1. $\mathrm{Tr}((1)_O) \Rightarrow \mathrm{Tr}((\text{‘p and not False ((1))’})_O)$ (for by definition (1): ‘p and not False((1))’)

[4] I consider contextualist objections to this line of reasoning in Section 2.4.1.

2. Tr((‘p and not False((1))’)$_O$) ⇒ (p and not False ((1))$_O$) (by one conditional (namely, Tr-release) of the Tr-principle)
3. (p and not False((1))$_O$) ⇒ not False((1)$_O$) (conjunction)
4. not False((1)$_O$) ⇒ Tr((1)$_0$) (as (1) is not gappy)
5. So Tr((1)$_O$) ⇔ (p and not False((1))$_O$)

This already seems like a good reason to believe that (1) is, so to speak, redundant: that it *only* says of itself that it is true, and that by saying that, it also says that *p*. In fact, I think that mere equivalence is not enough for identity of statements.[5] But we have more than a mere equivalence here: (i) the above reasoning is very simple; and (ii) provided that ‘p’ is chosen appropriately, any competent speaker who masters the concepts involved in the understanding of ‘Tr((1)$_O$)’ should understand those involved in that of ‘(p and not False((1))$_O$)’. Accordingly, no competent thinker could believe that (1) is true while not believing that p and (1) is not false. Nor could any such thinker believe that p and (1) is not false without believing that (1) is true.

- **ISP** (Identity of Statement Principle) On any occasion of use, if what the sentence (s) says is equivalent to what the sentence (t) says, and if no competent thinker can believe what the one sentence says without believing what the other sentence says, then (s) and (t) say the same thing.[6]

---

[5] Familiar counter-examples include cases of hyperintensionality (‘x is triangle’ is equivalent to ‘x is a trilatere’), failure of logical omniscience (‘1+1=2’ is equivalent to Fermat’s last theorem). More relevant to our point, if the Tr-principle is accepted, then any sentence

- (s): ‘p’

is equivalent to

- (t): Tr((s))

If (t) and (s) said the same thing, then (s) would say of itself that it is true. As a consequence, any sentence would be a truth-teller. Some philosophers have used considerations of this kind to solve the Liar paradox. For example, Mills [14] claims that every sentence attributes truth to the proposition it expresses. Prior, following Buridan and Albert of Saxony, has claimed that any sentence asserts its own truth. Accordingly, the Liar would entail a contradiction and be false. If the claim of this paper is correct, however, such a ‘solution’ would be too costly to deserve its name, for, as we shall see, it would imply that any sentence is a truth-teller so that its truth value on a given occasion of use can be changed virtually at will. But one should not accept the claim that (s) and (t) say the same thing anyway: a competent thinker could in general believe p without mastering the concept of truth or the concept of a sentence referring to itself, thus without believing that (s) is true.

Interestingly, Read [18] presents a solution to the Liar which is akin to the ones just mentioned but which is immune to the objections I have just raised against them. I discuss this solution in fn. 14.

[6] Notice that the condition of equivalence is not redundant with the condition on beliefs. I cannot rationally believe that (I *(de se)* believe that it is raining) without believing that (it is raining and I *(de se)* believe that it is raining), and arguably no one can. The two embedded sentences nonetheless say different things.

In virtue of **ISP**, (1) exactly says of itself that it is true:

- (1) | Tr((1))

More generally, we have the following redundancy principle:

- **1st RP** (1st Redundancy Principle) On any occasion of use, if a sentence says both that p and that it is not false, then that sentence just says of itself that it is true.

It might be thought that the claim that a sentence says that it is not false *and that p* and the claim that it *only says that it is true* and nothing more are in conflict. In fact, they are not. If I hear John say that truth is important, I can say the same, though *indirectly*, by saying that what John has just said is true. In the same way, if a sentence says both that *p* and that it is not false, then that sentence just says of itself that it is true, but by saying that, it ipso facto (and indirectly) says that *p*.

In the exact same way, on any occasion of use, Tr(2) ⇔ (r or not False ((2))) (for simplicity, we will omit the occasion of use):

1. not False((2)) ⇒ (r or not False((2))) (disjunction)
2. (r or not False(2)) ⇒ Tr('r or not False((2))') (by one conditional, Tr-Capture, of the Tr-principle)
3. Tr('r or not False((2))') ⇒ Tr((2)) (for by definition (2): 'r or not False ((2))')
4. Tr((2)) ⇒ not False ((2))
5. So Tr(2) ⇔ (r or not False((2)))

Just as before, no competent thinker can believe one of the two equivalents without believing the other one, so by **ISP**, (2) just says of itself that it is true.

- **2nd RP** (2nd Redundancy Principle) On any occasion of use, if a sentence says both that r or that it is not false, then that sentence just says of itself that it is true.

Accordingly, both (1) and (2) name truth-tellers: they name sentences which say of themselves exactly that they are true. Yet it is all too easy to choose 'p' and 'r' so that (1) and (2) have different truth values on any occasion of use. Indeed, if on any occasion of use 'p' is false, (1) which says that (p and not False(1)), will also be false. Similarly, if 'r' is true on any occasion, (2), which says that (r or not False(2)), will not be false. Therefore, provided that 'p' and 'r' are chosen this way, (1) and (2) will be two truth-tellers with different truth values. This means that truth-tellers cannot all have the same truth value.

One question that should naturally arise at this point is the following: in virtue of what do different truth-tellers have different truth values? What should make this question puzzling is that all truth-tellers just say of themselves that they are true, and nothing more. Certainly, (1) and (2) are spelled differently. But the mere graphical difference cannot be relevant here. Despite their graphical difference, (1) and (2) both say of themselves that they are true and nothing more. So what could explain their semantic difference? One might suggest that it is due to a difference in the occasions of use. More precisely, truth-tellers could be implicitly indexical. Such a contextualist (indexical contextualist, to be precise) view might seem promising as a neighboring claim has been invoked to solve semantic paradoxes of self-reference (see Section 2.4.1).

But it is not very plausible here. First, one could chose ‘p’ and ‘r’ so that swapping the context of use of (1) and (2) would not affect their respective truth values (take $p = \bot$ and $r = \top$). Second, the arguments above truly show that we can construct truth-tellers saying anything we want at any accessible space-time position. More precisely, we saw that *for any statements p and r*, we can construct a truth-teller the use of which makes a statement that implies that p and is implied by r. So if truth-tellers contained an indexical component, it would have to be quite an extraordinary one. In particular, it would have to behave in a fully unpredictable way. It would not have anything like a classical Kaplanian ‘character’ regimenting the way its context of use determines its content and its semantic value. We will see that the contextualist understanding of the Truth-Teller is indeed misguided. Truth-tellers are assessment-sensitive (relative) rather than use-sensitive (contextual).

2.3 The Truth-Teller Is Relative

In order to make this point, I will now show that the truth value of a token truth-teller *on a given occasion of use* depends on the *context in which it is assessed*, and that it depends on that context in such a way that it can be changed virtually at will.

Consider, for example, the sentence ‘Tr(($\star$))’ which I wrote in my notebook a long time ago, say on occasion V. I immediately stipulated that ($\star$) names that very sentence, so that ($\star$) is a truth-teller.

- **Stipulation 1**, made by me at time t_1 of the occasion V:
  - ($\star$): Tr(($\star$))

Now, suppose I want to try to make a new stipulation. Suppose I want to try to stipulate that ($\star$) not only names the sentence mentioned above, but that on the occasion of use V, it also says that p and that ($\star$) is not false. I can always *try* such a stipulation and write on my notebook:

- **Stipulation 2**, made by me later, at the time t_2:
  - ($\star$) names the token sentence ‘Tr(($\star$))’ above;
  - $(\star)_V \mid$ p and not False (($\star$)).

I can always attempt this, but, of course, there are strict conditions on the success of stipulations in a public language. Some stipulations fail to designate anything (consider ‘let’s call “oarf” the biggest prime number’, or more sadly, ‘Let’s call “John” someone who is my son and who will never be a drug addict’). Some stipulations might designate something but fail to designate anything determinate (consider ‘let’s define “cammals” this way: (i) cammals are mammals; and (ii) cats are cammals[7]'). Finally, some stipulations are only problematic in that they might create ambiguities and inconsistencies in the language (consider 'Let's call my little dog "Napoleon Bonaparte"'). I can always try to make such stipulations, but by doing so, I might run into trouble at one point or another. We say that a stipulation is *totally successful* when it does not *add* any problem of this kind to our language.

Now it might be hard to specify general conditions for the full success of a stipulation. But consider a case in which a name has already been defined by a stipulation and in which we want to make a new stipulation using this very name. A sufficient condition for this new stipulation to be fully successful is that it states the exact same thing as the old one. For example, if the term 'frenchlors' was defined as meaning French bachelors, nothing bars my stipulating that by 'frenchlors' I mean unmarried Frenchmen. Such a stipulation might be useless, but connectedly, it fully succeeds.

What about stipulation 2? If this stipulation says the same thing as stipulation 1, then it fully succeeds. Let us check. ($\star$) is used to name a sentence token that says of ($\star$) that it is true. My last stipulation has it name the very same token grapheme, and adds that on occasion V, this token says both that p and that it (($\star$)) is not false. But as we already saw (by **1st RP**), once I have stipulated that ($\star$) says that p and that ($\star$) is not false, saying (that p and that ($\star$) is not false) is the same as saying (that ($\star$) is true). Accordingly, once it has been made, my last stipulation states the same thing as the first one. So *once it has been made, it is bound to fully succeed.*

This means that for an arbitrary statement p, we can stipulate on time t_2 that what $(\star)_V$ said implies p. An analogous procedure allows to stipulate that what $(\star)_V$ said is implied by r:

- **Stipulation 3**, made by me at time t_3:
  - ($\star$) names the token sentence 'Tr(($\star$))' above;
  - $(\star)_V \mid$ r or not False(($\star$)).

Similarly (by **2nd RP**), such a stipulation is bound to fully succeed and it makes it the case that r implies $(\star)_V$. If we take r=$\neg$ p, then, whereas at t_2 $(\star)_V$ implied p, at t_3, $\neg$ p implies $(\star)_V$. Thus, we can have $(\star)_V$ say virtually anything we want. More precisely, for any statements p and r, a subject S considering $(\star)_V$ can have what it says imply q and be implied by r. This might seem problematic because two different subjects at a given time or one subject at two different times can make $(\star)_V$ have different truth values. But it is easy to see that this just entails a form of relativism. Let us call *context of assessment* for $(\star)_V$ the concrete situation in which $(\star)_V$ is assessed (such a situation should determine a subject assessing $(\star)_V$ and the time at which he assesses it, but it might include other things relevant for $(\star)_V$'s assessment). If (pace dialetheism) we should always avoid contradictions, the possibility of making the kind of stipulations that we made above implies that the truth value of $(\star)_V$ must depend on the context of its assessment. The truth of a truth-teller use is *relative*. We saw that there are two ways to interpret this.

[7] I owe this example to Williamson [22].

One can attribute the variations of the truth value with the context of assessment either (i) to a dependence of *the statement made* by the sentence use on the context of assessment, or (ii) to a dependence of the *truth value of the statement made* on the context of assessment, or (iii) to both kinds of dependences.[8] (iii) is not economical. According to (i), $(\star)_V$ makes different statements in different contexts of assessment. These statements have an absolute truth value but their existence is relative to some contexts of assessment. In other words, they are *semantically absolute* but *existentially relative*. In particular, they are short-lived, labile. According to (ii), $(\star)_V$ makes only one, eternal, *existentially absolute* statement, but it is a *semantically relative* statement: its truth value is labile; it depends on the context of its assessment. Whereas (1) is a form of *expressive relativism*, (2) is a form of *statement-relativism* (see Fig. 2.3). Statements being individuated by the way they contribute to distinguishing among both beliefs and assertions, the choice between (i) and (ii) depends on whether we would be inclined to say that someone who believes in/asserts the truth of $(\star)_V$ both in V and in V' truly believes/asserts the same thing in both contexts of assessment.[9] I am not sure what to think about this and I will try to remain neutral on this topic. I said a few times above that $(\star)$ says different things at different times. If (ii) is correct, this was speaking loosely. It would have been more accurate—although less convenient—to say that at different times, $(\star)$ entails different things or has different truth-conditions. In the same way, instead of saying that at time t_2, $(\star)_V$ says (among other things) that p, it would have been more accurate to say that at t_2 the truth of $(\star)_V$ entails that p. The task of rephrasing what precedes is left to the reader attracted to statement-relativism.

[8] There is actually a fourth part of this alternative. It might be the case that what $(\star)_V$ says is both absolute and constant; that is, that $(\star)$ always express the same thing, and that what it expresses has an absolute truth value. The arguments above would then show that $(\star)_V$ simultaneously makes a plurality of statements, some false, some true (actually, $(\star)_V$ would say virtually everything that can be said). The problem for such an *expressive pluralism* consists in accounting for the fact that the truth value of $(\star)_V$ can always vary with the context of assessment and turn out to be false in some contexts of assessment and true in others (notice that a pluralism like that of Read [18], which refuses **2nd RP** would avoid such a problem (see fn. 14)). One could try to account for such variations by distinguishing between the *statement made* by a sentence use, which is absolute and plural, and the *semantic content* of that sentence use relative to a context of assessment. Each context of assessment would select a salient semantic content from among the statements made by the sentence use, and this semantic content could turn out to be true. Cappelen [3] defends the idea that such a Pluralistic Content Relativism can account for most of the usual data cited by relativists. In the case at hand, a defender of this view must motivate the intuition that $(\star)_V$ says many things at a time (of assessment), and I am not sure how to do that.

[9] This is not to be confused with the question of whether if I believed/asserted that $(\star)_V$ is true at V, I should also believe it at V' and vice-versa. In general, and quite independently of relativism, there are often reasons reasons for not having the same attitude toward a single statement in different contexts of assessments. The relativity of the statement just adds one potential reason for such a change in attitude.

In any case, $(\star)_V$ is relative: its truth value is relative to the context of its assessment. More than that, it is, so to speak, *radically* relative.[10] Its truth value can be changed virtually at will, by stipulations of the kind we made earlier (stipulations 2 and 3). It can imply anything we want or be implied by anything we want. From now on, we will say that a sentence use is radically relative when its semantic value can likewise be changed at will.

Before we go on we should make a bit more precise the nature of the contexts of assessments for truth-tellers. A context of assessment for a sentence use is just a context in which this sentence use is assessed. We saw that it should include all the parameters which are relevant to determine the truth of that sentence use and that it can be represented by a n-tuple of such parameters. What allows to assign a truth value to a truth-teller use is a proof or an argument. If you take it, in a nominalist fashion, that all such proofs and arguments are real-world pieces of reasoning, then a context of assessment can be represented by the couple of a reasoner and a time. If on the other hand, you take it that truth value assignments can rely on immediate or timeless proofs or arguments, then contexts of assessment should include those abstract proofs or arguments which yield a truth value assignment. For the sake of simplicity, we have so far represented contexts of assessment as mere times, and we will often keep doing so. But it is important to keep the broader understanding of the contexts of assessment in mind when we state the general semantics of the sentence uses in scrutiny. Concerning the Truth-Teller, we should accordingly say the following:

- **Semantics of the Truth-Teller.**
  - Truth value assignments to truth-teller uses seem indeterminate because we only look for absolute-truth value assignments.
  - But truth-teller uses are relative: a truth-teller use is (i) true in a context of assessment in which it is proved to be true, (ii) false in a context of assessment in which it is proved to be false and (iii) indeterminate otherwise.
  - Contexts of both type (i) and type (ii) exist or can be constructed.

2.4 Other Pathologies of Self-Reference

2.4.1 The Liar

The Truth-Teller is not usually considered a true paradox. Like the Liar, it exhibits some kind of semantic deficiency which seems partly linked to its self-referential character (in the words of Kripke [10], it is ‘ungrounded’). But it seems only indeterminate, whereas paradoxical sentence uses seem to entail inconsistencies. In any case, we might suspect that the same mechanisms are at work in the various

[10] This use of ‘radically relative’ should not be confused with Recanati’s [19] use of the term. Recanati uses the moderate vs. radical relativism distinction to contrast what we called nonindexical relativism (his ‘moderate relativism’) with relativism (his ‘radical relativism’).

pathologies of self-reference, so it would be nice if we could propose a unified resolution of these pathologies.

The relativist approach to the Truth-Teller provides precisely such a unified resolution. I will not go into the details here. What precedes has already demonstrated the need to postulate a form of radical relativity. Thus, I will not try to show that the Liar also *necessitates* such a hypothesis. I will simply show how radical relativity can help solve the Liar.

Consider the type sentence 'not Tr ((L))'. I can call (L) a token of such a sentence written on this paper and used on occasion U, at time t_L:

- (L): 'not Tr((L))'

(L) is a (strengthened) Liar. It is problematic for the following reason. Suppose that

- (3): Tr((L))

then by definition

- (4): Tr('not Tr((L))')

and by Tr-release

- (5): not Tr((L))

This contradicts our initial hypothesis. But if not Tr((L)) (that is, if (5) is true) then, by Tr-capture

- (6): Tr('not Tr((L))')

By definition of (L) again,

- (7): Tr((L))

so (3), and not (5), should be true, etc.

The classical solution to the Liar consists in saying that (L) is neither true nor false. This can mean either that it has a third, gappy, truth value, or that it has no truth value at all because it fails to make a statement.

This solution runs up against the following 'revenge problem'. As (L) is neither true nor false, it is not true. But is this not just what (L) says? So (L) would be true after all. In other words, (5) must be true, but *(5) and (L) say the same thing*. So (L) must be true, and the paradox looms again.

One promising answer to this revenge problem consists in denying that (5) and (L) really say the same thing at the same time. This is the so-called '*contextualist approach*' (it is actually a form of indexical contextualism) to the semantic paradoxes. This might seem counterintuitive. After all, both (5) and (L) just say of (L) that it is true. How could they fail to say the same thing? The heaviest burden for the contextualist consists precisely in finding a plausible answer to this question. There are many different contextualist approaches. Some claim that the truth predicate is indexical, so that the meaning or extension of the predicate used in (5) and (L) is different [2, 5]. Others say that none of the constituents of (L) and (5) are indexical, but that (L), because of its paradoxical nature, fails to make a statement, whereas (5) makes a genuine (and indeed true) statement [7]. Still others claim that, because of its self-referential character, (L), but not (5), implies its own falsity, so whereas (L) is false, (5) is true [17, pp. 208, 220–21]. All the same, I cannot help but find ad hoc the claim that (L) and (5) differ in meaning or extension.

The notion of relative statements can help us find a solution to the strengthened liar that keeps something of the intuitions behind the contextualist approach, without having to claim that (L) and (5) differ in meaning or extension. According to this solution, $(L)_U$ might really say something, but it must be a relative, and indeed radically relative, sentence use. More precisely, no matter what (L) said when I used it on this paper, that is, on occasion U, at time t_L, and even if it said nothing at all, we can assert the following. When a little bit later on that paper, at time $t_{(5)}$, I used (5), it became true that (L) states something on occasion U, and that what it states is true. When I realized, at time t_7 when I used (7), that (7) is true, it *then* became true that what (L) states on occasion U is false, and accordingly that what (3) and (7) stated when they were used is now true, etc. Writing down a new sentence like (5) or (7) and using it changes the context of assessment of $(L)_U$, and thus changes its truth value. The dynamic process of stating the argument to the effect that $(L)_U$ is paradoxical changes the context of assessment and thus the truth value of $(L)_U$. And this allows us to avoid (synchronic) contradictions (see Fig. 2.2).

What distinguishes liar sentence uses from truth-teller sentence uses according to this account is not that the latter can say or imply virtually anything and have any truth value whereas the former cannot be assigned any truth value. Liar sentence uses could actually say or imply many things, too.[11] What makes them special is that they cannot have a stable truth value. More precisely, if we say that a sentence use is *convergent* when there is a moment of assessment from which its truth value does not vary, then liar sentence uses cannot be convergent. Whereas the truth value

time of assessment	$t_{(5)}$	$t_{(7)}$	$t_{(9)}$	$t_{(11)}$	...
truth value	false	true	false	true	...

Fig. 2.2 The truth value of the use of (L) at occasion U relative to the time at which I assess it

[11] We could treat the Liar as we treated the truth-teller, showing that for any sentence p, one can make fully successful stipulations to the effect that a liar sentence says that p, even though for the usual problems, such a stipulation could not rationally be sustained very long.

Let us outline the argument roughly. In virtue of the Tr-schema, not Tr((L)) ⇔ Tr('not Tr((L))'). But of course, if I am to believe that not Tr((L), I must master the concept of truth and that of a sentence referring to itself. Accordingly, I, and this holds for any competent thinker, cannot believe that not Tr((L)) without believing that Tr('not Tr((L))'), and cannot believe that Tr('not Tr((L))') without believing that not Tr((L)).

So by **ISP** (L) says exactly that it is true. We can then proceed exactly like in the case of the truth-teller.

of a truth-teller can converge (provided that from a given moment on no one changes it) the truth value of a liar sentence use is bound to change forever with the time of assessment.

This makes liar sentences quite *sui-generis*.

- Regular, ordinary, well-behaved sentence uses have an absolute truth value.
- If MacFarlane and other contemporary relativists are right, some sentence uses dealing with future contingents and matters of taste and some sentence uses ascribing knowledge or justification are relative [11, 12]. There are, however, limits on the relativity of such sentence uses. One of these sentence uses could only say or imply finitely many things, and its truth value would be convergent. Those sentence uses, in other words, would only be moderately (as opposed to radically) relative and would not truly be pathological.
- By contrast, truth-teller uses are radically relative. While they may fail to be convergent, they do not have to.
- Liar sentence uses are also radically relative, but they cannot be convergent.

Again it should be noticed that this solution does not depend on the claim that proofs are extended in time. If you admit timeless proofs, you should include those proofs in your contexts of assessment. $(L)_U$ would then be relative not to the time of assessment but to the proof with which it is assessed at that time. It would be false relative to the proof we wrote between $t_{(1)}$ and $t_{(5)}$, true relative to the proof we wrote between $t_{(5)}$ and $t_{(7)}$, false relative to the proof we wrote between $t_{(7)}$ and $t_{(9)}$, etc. This allows to avoid contradictions relative to a given context of assessment. The derivation of a contradiction is indeed blocked by the fact that one cannot derive the truth of a sentence use in a context of assessment from its truth in another context of assessment. With such a broadened understanding of the contexts of assessment, the distinction between the semantics of liars and that truth-tellers is however a bit harder to pin down. One might want to say that for a truth-teller use, for each context of assessment in which it is true there is another context of assessment in which it is false. But that seems to be true of truth-tellers as well if contexts of assessments are defined as abstract proofs in a timeless world and not as real-world pieces of reasoning. We can generalize the concept of convergence and make the difference between liars and truth-tellers explicit by invoking a rational subject. In the case of liar sentence uses, but not in the case of truth-tellers, for each context of assessment there is always a context of assessment *that a subject should consider if she is rational* in which the sentence use has a different truth-value. Liar sentences, but not truth-tellers, are in that sense convergent.

- **Semantics of the Liar:**
  - Truth value assignments to liar uses seem inconsistent because we only look for absolute-truth value assignments.
  - But liar uses are relative. A liar use is (i) true in a context of assessment in which it is proved to be true, (ii) false in a context of assessment in which it is proved to be false, (iii) and indeterminate otherwise.

– Moreover, a rational subject should not consider a context of type (iii) without subsequently considering a context of type (i). And she should not consider a context of type (i) or (ii) without considering one of the other type next (non convergence).

This relativist solution to the Liar and the Truth-Teller has a cost. It claims that the truth value of some sentence uses is relative. But this cost should not be overestimated. First, it allows us to attribute semantic-values to the pathological sentences while conserving classical logic and the naive (unrestricted) Tr-schema. Second, we showed that if we wanted to attribute semantic-values to the pathological sentences and to conserve classical logic and the naive Tr-schema, we *had* to buy relativism: we came to this solution deductively rather than abductively. Third, relativism is not as problematic as might be thought. Others have shown how to make sense of relative truth [12]; this is not the place to rehearse their arguments.[12] Fourth, and more importantly, relativism is not especially problematic when it is *relativism about pathological sentences* such as the ones we considered. The classical concern about relativism is that there seems to be a constitutive link between assertions, beliefs and truth, and that the relativity of truth seems to threaten this link: truth is something assertions aim at, but the aim of assertions is absolute, so truth cannot be relative. MacFarlane [12] proposes reformulations of the link between truth and assertion that neatly accommodate relative truth (see also [11, pp. 333–35]). But such a reformulation is not even required in our case, because the pathological sentences we considered are such that no one should assert or believe what they say anyway. One should not play the game of assertion with them, and indeed, it is usually only by mistake that one ever starts playing such a game with them.

2.4.2 Other Semantic Pathologies

As far as I can tell, all semantic pathologies of self-reference can be solved in an analogous fashion. Consider for example the famously resistant Open-Pair:

- (8): the neighbouring sentence is false
- (9): the neighbouring sentence is false

---

[12] Some seem to believe that relativism threatens the Tr-schema and that, because our argument for relativism relied on the Tr-schema, it is self-defeating. This is not the case. We can easily reformulate the version of the Tr-schema we relied on so as to accommodate relativism. The following would work:

- If (x) names a sentence, and $(x)_U$ names the use of (x) is occasion of use U. If $(x)_U$ sates that P in the context of assessment A,
  - $(x)_U$ is true in A iff P

  We can consider the reference to the context of assessment as implicit in the first pages of this paper.

This pair of sentences is paradoxical, or at least pathological, because neither (8) nor (9) can be gappy (being neither false nor true would make them false) and because they cannot have the same truth value. But for reasons of symmetry, there seems to be no grounds not to attribute the same truth value to them, hence the paradox. The notion of relative sentence use allows us to solve that paradox. In the same way that a single time-slice of a four-dimensional cyclical worm can hide its overall symmetry, neglecting the variations between the open pair's truth value assignments and the time, or more generally the context, of its assessment might hide their perfect symmetry. The variations of the truth values with the time of assessment add a new semantic dimension, a dimension which might reveal an otherwise hidden symmetry.

The paradox indeed disappears once we allow ourselves relative sentence-uses. Using the now-familiar kind of argument, on any occasion of use, *both* (8) and (9) can imply or be implied by virtually anything we want at any time of assessment. Even better, we can assign them at will any pair of truth values compatible with what they say. As they say of each other that they are false, we will have to assign them opposite truth values at any time of assessment. If the consideration of symmetry truly is telling, we will have to say that if, at a given context of assessment t, (8) says something false on the occasion of use V, and (9) says something true on that very occasion, there is a context of assessment t' at which this assignment of truth value is reversed, and this reversal balances the initial asymmetry. The asymmetry seemingly required by the contradictory contents of (8) and (9) was generated by a static point of view. Once we introduce the dynamic point of view connected with relativism, the need for asymmetry, and the paradox it creates, both vanish.

As before, this solution is not restricted to temporal contexts of assessments:

- **Semantics of the Open-Pair:**
  - Truth value assignments to open-pair uses seem arbitrarily asymmetric because we only look for absolute-truth value assignments.
  - But open-pair uses are relative. An open-pair use is (i) (true,false) in a context of assessment in which it is proved to be so, (ii) (false,true) in a context of assessment in which it is proved to be so, (iii) and indeterminate otherwise.
  - Moreover, a rational subject should not consider a context of type (iii) without subsequently considering a context of type (ii) or (iii). And she should not consider one of those last two context types without considering the other one next (non convergence).

2.4.3 Immunity to Revenge Problems

Most, if not all, traditional solutions to the paradoxes of self-reference are famously thwarted by a revenge problem: the semantic notions they introduce to solve the paradoxes allow for the production of new paradoxes that exactly mimic the old, 'solved' ones, but which cannot be solved in the same way. For example, assigning a third, gappy, truth value to 'this sentence is false' allows us to build a new, strengthened, paradox 'this sentence is not true, that is, either false or gappy' which, we argued, cannot be gappy. As far as I can tell, the solution in terms of relative truth presented here is *immune to such revenge problems*. To see this, consider some of the most likely candidates (other likely candidates can be treated in the same way):

- (10): (10) is not true in the context of assessment C.
- (11): in all contexts of assessment (11) is false.
- (12): in all contexts of assessment, (12) is true.

One could think that in order to deal with a given use of (10), and avoid a contradiction at the context of assessment C, we must introduce the notion of truth which is so to speak 'doubly relative' ([truth at a context of assessment] relative to the context in which this relative truth is itself assessed) or enrich the contexts of assessment to accommodate a wider source of relativity. No such complication is actually needed. The basic point is that the paradoxical or pathological sentences are not self-assessing: their assessment is not logically immediate, they all need some kind of proof or argument in order to be assessed and assigned a truth value. (One might even suspect that this has to be so. If a sentence were self-assessing it would *trivially* admit a consistent and determinate truth value assignment or *trivially* fail to admit any such assignment. In either case we wouldn't be tempted to say that it is problematic and it would not be paradoxical or even pathological in the relevant sense.) So one cannot assess the [truth value of a use of (10) in context C] in that very context C, but only in a modified context C'. And there cannot be a proof of both the truth and non truth of (10) at C. No revenge problem is looming.

(11) is paradoxical, like the regular liar, but it can be solved in the same way as the latter: it can be argued that it is true relative to some contexts of assessment and false relative to others. Indeed, assuming that when a sentence says that is not true *simpliciter*, it implicitly says that it is not true relative to all contexts of assessment, (11) is nothing but a fancy reformulation of the regular liar-sentence. (12) is likewise a fancy reformulation of the regular truth-teller and it can be treated in the same way as the latter.

At that point one might wonder if there is not, however, a problem with the claim that (11) and (12) are strictly relative (true relative to some contexts and false relative to others). Both (11) and (12) say of themselves that they have the same truth value relative to all contexts, so one might suspect that they cannot be strictly relative: if they were strictly relative, they would be false relative to all contexts of assessment, and thus *not* strictly relative. This reasoning is however misguided. Once the hypothesis of their relative truth is envisioned, it cannot be proved that those sentences are strictly relative then they are *absolutely false*. It can only be proved that if they are strictly relative, then are false *relative to the context in which this proof is provided*.

Revenge problems do not arise when the new, paradox-solving, semantic status allows to formulate new sentences which are likewise paradoxical. They arise when it allows to formulate new sentences *which cannot have this new semantic status* but are likewise paradoxical nonetheless. Now there is a fundamental reason why the relativist solution is immune to revenge problems. It is that there is no way to use the notion of relative truth in order to produce sentence uses that cannot be relative.

In particular, trying to explicitly mention the context of assessment in a sentence so that it cannot be relative betrays a misunderstanding of the difference between indexical contextualism and relativism. According to indexical contextualism, some sentences have open argument-places whose value is provided by the context of use. Those sentences are semantically incomplete because they cannot be evaluated unless this value is provided. We can obtain a-contextual sentences by saturating those incomplete sentences with the explicit mention of the context in which they are used. Thus, whereas 'I am a philosopher' is contextual 'Alexandre Billon is a philosopher' is not.

According to relativism, on the other hand, some sentences are relative to a context of assessment even though they are semantically complete: they are ready to be evaluated, all we need to evaluate them being provided by the process of evaluation itself. There is thus no sense in trying to 'saturate' such sentences. We can construct, as we did, sentences of the form 's relative to C' or "s' is true relative to C' (where 's' is substituted with a relative sentence), but this will not guarantee that the resulting sentence is not relative: its truth-value might well vary with the contextual parameters involved in the process of its evaluation.

2.5 Dissolutions, Cassations and Resolutions

Roughly speaking, if semantical paradoxes are paradoxical, it is because, even though they seem to be perfectly grammatical, declarative sentences, it appears that they cannot have a semantic value at the risk of being contradictory. More broadly, if semantical pathologies are pathological, it is because, even though they seem to be perfectly grammatical declarative sentences, they do not seem to have a *determinate* semantic value. Solutions to these paradoxes try to accommodate conflicting intuitions in one way or another. I can distinguish three approaches to the semantical pathologies of self-reference.[13]

1. *Dissolution.* One can first deny that there is anything pathological with the sentences at hand, and show that if we are more cautious in spelling out their properties, problematic sentences actually behave like ordinary, benign sentences. Prior, Mills [14], Bardwardine and Read [18] endorse such an approach.
2. *Cassation.* One can, on the contrary, acknowledge the peculiarity of the problematic sentences, but deny that those sentences say anything at all. They simply fail to make a statement. This is the approach favoured by Goldstein [7].
3. *Resolution.* A third approach also acknowledges the peculiarity of the problematic sentences. Instead of denying that they express anything at all, partisans of such an approach typically insist that they express some peculiar or unusual statements. Some will, for example, say that the Liar, although it really says

[13] This list does not claim to be exhaustive. Thomas Bradwardine famously listed 9 different views on the *insolubilia*. I suspect that some of them would not belong on my list.

something, says something that is neither true nor false [10]. Others will claim that it says something that is both true *and* false [15], etc.

As should be clear by now, the solution I proposed is of the third kind. It resolves the paradox by saying that the problematic sentences do have an absolute truth-value but only a relative one. I favour this third, or resolutionist approach because I believe the first approach, or at least what I know of the solutions which come under it, to be a nonstarter.[14] I favour it, also, because, although I have sympathies for the second, cassationist approach, I believe that the resolutionist approach is (i) more fruitful; and (ii) that it is conservative enough to accommodate most of the cassationist insights. Let me give a very close analogy from algebra. As is well-known, there was once a problem with equations like

- (e) $x^2 = -1$

It seemed that no number could satisfy such an equation because the square of every number is positive. Still, in some fields of physics or mathematics, one could end up looking for a number satisfying (e). Cassationists of that time maintained that such a number did not exist. (e) simply failed to pick a number, and that was all there was to say. Resolutionists hypothesized, on the other hand, that (e) might actually pick a number, but a number of a special kind, call it imaginary, and different from the usual *real* numbers. The resolutionist approach proved very fruitful over the years. But it also showed that it could accommodate the cassationist intuitions very well. In fact, the terminology of *real* and *imaginary* numbers seems tailored for that: usual, well-behaved numbers are *real* numbers, and (e) fails to pick a *real* number. Considerations of the same kind apply in our case. It is very easy to rephrase the

---

[14] I criticized the 'solutions' put forward by Mill and Prior in fn. 5. Read proposes a solution to the paradoxes which, as he pointed out to me, escapes those criticisms and seems to be in a position to dissolve the problems posed by (1) and (2). His solution relies on what he calls a meaning or signification pluralism:

- **MP**: a sentence signifies more than one thing; it signifies not only what its terms explicitly say, but also all the consequences of what it signifies.

Read uses this meaning pluralism to defend a principle connecting truth and signification aimed at replacing, or at specifying the Tr-schema:

- **TS**: a sentence use is true iff everything it signifies (as opposed to what its terms explicitly say) is true.

The main point is that this principle does not in general yield Tr-capture. It only yields Tr-release. So according to Read, our argument for the **1st RP** would be sound, but not that for the **2nd RP**. We could only show that some truth-tellers are false (not that some are true) and that it is always possible to make a truth-teller turn false by a mere stipulation. But the best explanation for this last fact would be that all truth-tellers are false. In the same way, the truth of the liars would entail their falsity, but without Tr-capture, their falsity would not entail their truth. As a consequence, liars would just be false.

I have two concerns with this elegant proposal. First, I am not totally convinced that it is not ad hoc. Second, I am not sure that it can account for the intuitive difference in meaning between liars and truth-tellers and for that between liars and ordinary contradictions of the form 'p and not p'.

cassationist's claim in a way that will make it compatible with the resolutionist's approach: instead of saying that pathological sentence uses fail to make a statement, we can say that, contrary to well-behaved, ordinary sentence uses, they are radically relative, and that, accordingly, they fail to make a (unique) ordinary, well-behaved statement. They fail to pick a (unique) statement that is not existentially or semantically (radically) relative.

I said of the resolutionnist approach that it is not only conservative, it is also more fruitful than the cassationnist approach. So let me conclude with what I take to be one of the most promising (if still slightly speculative) outcomes of the solution I propose.

There is a long tradition in philosophy of distinguishing between two ways in which a feature of our mental life can be intrinsically meaningful. Some are intrinsically meaningful because they are intentional. This is the case, for example, of beliefs, fears, perceptual experiences, etc. They intrinsically mean something to their subject because, just in virtue of having them, the latter is assessable for truth or accuracy. There are mental features, however, which seem to be intrinsically meaningful for their subject but which do not seem to be intentional. Some moods (nervousness, depression, elation and undirected anxiety), drives, mere feelings, or raw sensations do not seem to be about anything, and their subject does not seem to be assessable for a semantic value in virtue of having them. Yet they are not like, say, single neuron firings, tables or chairs. They intrinsically mean something to their subject. Phenomenologists, who like to put meaning talk in terms of appearance and manifestation would say that those states make something manifest to their subject. One might also say, more simply, that they are *felt* by him. We can call *meaning dualism* the claim to the effect that such 'mere feelings' constitute a distinct form of meaningfulness which is irreducible to intentionality.

Although such a meaning dualism is not obvious, and has been denied, it is quite popular and I would say, quite attractive (see [20, chap. 1] for a recent defense, and [4] for a criticism). It is, however, usually restricted to the mind, and it is not clear how an analogous claim concerning language could be defended or even articulated.[15] My solution is fruitful because it allows to do both things, and to connect some worries in the philosophy of mind with traditional problems in the philosophy of language. Let me give a brief outline.

What distinguishes mere feelings from intentional mental states is that intentional features are *intrinsically* assessable for a semantic value (or, if you prefer, their

[15] An interesting suggestion would be that some sentences are exclusively expressive (as opposed to, say, assertive or descriptive) and that those are the linguistic analogue of mere feelings. A claim along this line is endorsed by classical expressivists in ethics [1]. As speech-acts expressing mere feelings seem continuous with speech-acts expressing intentional states it is not clear that such a suggestion is compatible with the plausible claim to the effect that the semantic properties of speech-acts reflect those of the mental state they express.

subject is so assessable just in virtue of them) whereas mere feelings need to be provided an interpretation in order to be assessable for a semantic value. They are only assessable *relative to an interpretation*. In that respect, mere feelings resemble single neuron firings table and chairs. Both of them are only assessable relative to an interpretation, and both of them can be interpreted and reinterpreted at will, by mere stipulations. What distinguishes mere feelings from single neuron firings, etc., however, is that they are intrinsically meaningful to the subject and seem to call for an interpretation. We saw that a sentence is radically relative just in case it can be interpreted and reinterpreted at will by mere stipulations, just in case, that is, the context of its assessment can be considered as providing an interpretation for it. We could say, accordingly, that feelings are to be distinguished from single neuron firings on the one hand and from intentional states on the other hand by the fact that they are *intrinsically meaningful but radically relative*. Meaningful but radically relative sentences like those (interpreted) pathological sentences I dealt with in this paper would thus be the linguistic analogue of mere feelings. They should be distinguished from uninterpreted sentences and regular, non pathological, sentences just like feelings are distinguished from single neuron firings and intentional states (Fig. 2.3).

I come from the philosophy of mind and I am definitely a meaning dualist concerning the mind. I believe, in particular, that such features like consciousness and subjectivity (the typical me-ishness of my conscious states) that are reputed to be irreducibly non intentional really are non intentional, and can only be understood in terms of mere feelings. I became interested in the semantic pathologies of because some philosophers have thought they could circumvent most of the difficulties that afflict classical intentionalist analyses of consciousness and subjectivity by putting forward a *self-referential* intentionalist analysis (see [23]). Along with Marie Guillot, we criticized those self-referential intentionalist theories. We argued that those theories fail because the self-referential mental states they posit must be *pathologically* self-referential [8]. I submit that ultimately, they fail because some properties of consciousness and subjectivity cannot be explained by intentional features. They can only be explained by 'mere feelings', by radically relative, but meaningful, mental features.

	Meaningless	**Meaningful**
Radically relative	uninterpreted sentences; single neuron firings, tables, etc.	liars, truth-tellers and other radically relative sentences; mere feelings, etc.
Not radically relative		non pathological sentences; intentional states.

Fig. 2.3 Meaning dualism for the mind and language

Acknowledgements Many thanks to Shahid Rahman, Stephen Read, Tero Tulenheimo, Mathieu Fontaine, Laurent Keiff, the audience at the 2009 SOPHA conference for useful discussions and to Marie Guillot for a priceless collaboration on related topics.

References

1. Ayer, A. 1958. *Language, Truth and Logic (1936)*. London: Victory Gollancz.
2. Burge, T. 1979. "Semantical Paradox." *Journal of Philosophy* 76:169–98.
3. Cappelen, H. 2008. "Content Relativism and Semantic Blindness." In *Relative Truth*, edited by M. Garciá-Carpintero and M. Kölbel, 265–86. Oxford: Oxford University Press.
4. Crane, T. 1998. "Intentionality as the Mark of the Mental." *Royal Institute of Philosophy Supplement* 43:229–52.
5. Gaifman, H. 1992. "Pointers to Truth." *The Journal of Philosophy* 89(5):223–61.
6. Garciá-Carpintero, M., and M. Kölbel, eds. 2008. *Relative Truth*. Oxford: Oxford University Press.
7. Goldstein, L. 2000. "A Unified Solution to Some Paradoxes." In *Proceedings of the Aristotelian Society, New Series*. The Aristotelian Society, Blackwell Publishing.
8. Guillot, M., and A. Billon. 2007. "Could Subjective Access to One's Thoughts Be a Matter of Self-Reference?" Talk at ENS Lyon (May 5, 2007), paper in preparation.
9. Kölbel, M. 2008. "Introduction: Motivations for Relativism." In *Relative Truth*, edited by M. Garciá-Carpintero and M. Kölbel. Oxford: Oxford University Press, 1–40.
10. Kripke, S. 1975. "Outline of a Theory of Truth." *Journal of philosophy* 72(19):690–716.
11. MacFarlane, J. 2003. "Future Contingents and Relative Truth." *The Philosophical Quarterly* 53(212):321–36.
12. MacFarlane, J. 2005. "Making Sense of Relative Truth." In *Proceedings of the Aristotelian Society, New Series*, vol. 105, 321–39. The Aristotelian Society, Blackwell Publishing.
13. MacFarlane, J. 2009. "Non Indexical Contextualism." *Synthese* 166(2):231–50.
14. Mills, E. 1998. "A Simple Solution to the Liar." *Philosophical Studies* 89(2–3):197–212.
15. Priest, G. 2006. *In Contradiction* (2 Edition). New York, NY: Oxford University Press.
16. Rahman, S., T. Tulenheimo, and E. Genot, eds. 2008. *Unity, Truth and the Liar,* vol. 8 of *Logic, Epistemology, and the Unity of Science*. Dordrecht: Springer.
17. Read, S. 2008a. "Further Thoughts on Tarski's T-scheme and the Liar." In *Unity, Truth and the Liar,* vol. 8 of *Logic, Epistemology, and the Unity of Science,* edited by S. Rahman, T. Tulenheimo, and E. Genot, chapter 13, 205–25. Dordrecht: Springer.
18. Read, S. 2008b. "The Truth Schema and the Liar." In *Unity, Truth and the Liar,* vol. 8 of *Logic, Epistemology, and the Unity of Science,* edited by S. Rahman, T. Tulenheimo, and E. Genot, chapter 1, 3–18. Dordrecht: Springer.
19. Recanati, F. 2007. *Perspectival Thought: A Plea for (Moderate) Relativism*. Oxford: Clarendon Press.
20. Searle, J. 1983. *Intentionality: An Essay in the Philosophy of Mind.* Cambridge, MA: Cambridge University Press.
21. Smith, J. W. 1984. "A Simple Solution to Mortensen and Priest's Truth Teller Paradox." *Logique et Analyse* 27:217–20.
22. Williamson, T. 2000. "Review of 'blindspots' by Roy Sorensen." *Mind* 99(393):137–40.
23. Williford, K., and U. Kriegel, eds. 2006. *The Self-Representational Approaches to Consciousness*. Cambridge, MA: MIT Press/Bradford Books.
24. Woodbridge, J. A. 2004. "A Neglected Dimension of Semantic Pathology." In *Logica Yearbook*, 277–92. Prague: Filosofia, Institute of Philosophy, Academy of Sciences of the Czech Republic.

Chapter 3
Which Logic for the Radical Anti-realist?

Denis Bonnay and Mikaël Cozic

3.1 Introduction

One of the most striking outcomes of the controversy between semantic realism and semantic anti-realism concerns logic: it is widely held that an anti-realist position should result in a *revisionist* attitude with respect to classical logic. More precisely, according to the seminal contributions of M. Dummett, a coherent anti-realist should prefer intuitionistic logic to classical logic. Therefore, one would expect that different forms of anti-realism would result in different forms of revisionism. Recently, it has been argued by J. Dubucs and M. Marion [4, 5] that an anti-realism more radical than the usual one justifies another logic than the intuitionistic, namely *linear logic* [9]. If one calls "moderate" the version of anti-realism advocated by Dummett, the landscape is going to be the following:

realism	⇒	classical logic
moderate anti-realism	⇒	intuitionistic logic
radical anti-realism	⇒	linear logic

The aim of this paper is not to take a position on the left-hand side of the tabular. We shall take position neither in the realism/anti-realism debate nor in the moderate/radical anti-realism debate. Our focus is the idea that a strengthening of the moderate anti-realist's basic insights leads to linear logic rather than to intuitionistic logic. We shall ask whether there is a path from the bottom-left cell to the bottom-right cell that is parallel to the usual path that goes from the middle-left cell to the middle-right one.

D. Bonnay (✉)
Université Paris Ouest-Nanterre, Ireph & IHPST, Paris, France
e-mail: denis.bonnay@ens.fr

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_3,

We will proceed as follows. In Section 3.1, we give a rough reconstruction of anti-realism's basic tenets and of the substructural revisionism of radical anti-realists. We then scrutinize both the consequences of committing oneself to substructural revisionism and the principles that could back up this commitment. In Section 3.2, we argue that, because of the splitting of connectives, it is not easy to live without structural rules. Therefore the justification for such a shift has to be pretty firm. But in Section 3.3, we show that there is currently no satisfactory foundation for substructural revisionism. In Section 3.4, nonetheless, we briefly sketch a possible, game-theoretic, way to achieve such a foundation.

3.2 From Anti-realism to Substructural Logic

3.2.1 Moderate Anti-realism

We shall first reconstruct briefly the position of *moderate* anti-realism. Though we do not want to enter into an exegetical discussion, the view we present here could be called Dummettian anti-realism. We take moderate anti-realism to consist in two basic components: the anti-realist component per se and the revisionist component.

Moderate anti-realism starts with a rejection of truth-conditional semantics. According to truth-conditional semantics, the meaning of a declarative sentence S is the condition under which it is true—and to grasp the meaning of a sentence S is to grasp its truth-conditions. Furthermore, the truth or falsity of S is independent of our means of knowing it[1]: nothing precludes that the conditions under which S is true cannot be recognized as such when they obtain. To put it another way, S could be true even though it is not possible to know that it is true. According to the realist, truth in not epistemically constrained.

The anti-realist rejects precisely this lack of epistemic constraint: if a sentence S is true, then it should be possible to recognize that it is true. (This is the so-called "Knowability Principle".) There are two main arguments in favor of the Knowability Principle: if knowledge of meaning is to be analyzed as knowledge of truth-conditions, one has to be able to gain such knowledge (this is the *learnability argument*) and to manifest that one possesses it (this is the *manifestability argument*). As long as truth-conditions are recognition-transcendent, knowledge of truth-conditions does not satisfy these requirements. The realist thus fails to account for our mastery of language.[2]

---

[1] See [6, p. 146], "Realism": "Realism I characterize as the belief that statements of the disputed class possess an objective truth-value, independently of our means of knowing it [. . .]"

[2] In the case of mathematical discourse, see [6]: "If to know the meaning of a mathematical statement is to grasp its use; if we learn the meaning by learning its use, and our knowledge of its meaning is a knowledge which we must be capable of manifesting by the use we make of it: then the notion of truth, considered as a feature which each mathematical statement determinately possesses or determinately lacks, independently of ours means of recognizing its truth-value, cannot be the central notion for a theory of the meanings of mathematical statements [. . .]" (p. 225).

As a consequence, the anti-realist rejects the notion of truth-conditions as an adequate basis for a theory of meaning, and puts forward as an alternative the "conditions under which we acknowledge the statement as conclusively established" [6, p. 226], or, as it is sometimes put, assertibility-conditions, i.e. the conditions under which one is justified to assert the sentence. When it comes to mathematical discourse, one is justified to assert a sentence just in case one has a proof of that sentence. Therefore, the meaning of a mathematical sentence consists in its provability conditions (as opposed to its mysterious recognition-transcendent truth-conditions).

This is it for our reminder of the basic tenets of moderate anti-realism. What we want to stress now is that (and how) *these tenets lead to logical revisionism*: they give us strong reasons to reject classical logic and, at least as far as mathematical discourse is concerned, to endorse intuitionistic logic. As a matter of fact, the path from anti-realism to logical revisionism is not as clear as one might wish. Actually, we think that there are *two ways* from anti-realism to logical revisionism: one that goes directly from the rejection of realism to the rejection of the law of excluded middle; and one that goes through proof-theoretic arguments from the endorsement of anti-realism to a thorough justification of intuitionistic logic[3] We shall call the first way *"high-level revisionism"* and the second *"low-level revisionism"*.

Let us elaborate on this distinction, which will play an important role in our discussion of radical anti-realism. *High-level revisionism* consists in rejecting the excluded middle on account of the Knowability Principle. More precisely, for Dummett, the rejection of the law of excluded middle (LEM for short) stems from the rejection of the principle of bivalence according to which every (meaningful, non-vague and non-ambiguous) declarative sentence is determinately true or false. Bivalence is not equivalent to LEM, but, as Dummett puts it, "once we have lost any reason to assume every statement to be either true or false, we have no reason, either, to maintain the law of excluded middle" [7, p. 9].

The exact argument for the rejection of bivalence is a matter of controversy.[4] J. Salerno has convincingly argued that Dummett's and Wright's (see [20, p. 43]) arguments are unsound, but he has also proposed an amended version. His point is that the following three are incompatible:

(i) It is known that LEM holds.
(ii) The Knowability Principle is known: we know that for all A, if A is true, then it is possible to prove that A.

[3] See [14] for a closely related presentation of the anti-realist case for revisionism: what we call "high-level revisionism" corresponds roughly to what Read calls the "Linguistic Argument" and what we call "low-level revisionism" corresponds to what he calls the "Logical Argument". The distinction is implicit in other places: e.g. in Tennant's [19], chapters 6–7 focus on "high-level revisionism", whereas chapter 10 focuses on "low-level revisionism".

[4] The question whether Dummettian anti-realism succeeds in vindicating logical revisionism has been (and is) much disputed. See C. Wright, "Anti-realism and revisionism" in [2, 3, 18–20]. In particular, [19] argues that Dummett's manifestation argument, even if it is an "attempted reductio of the principle of bivalence", "*in so far as it is directed against bivalence*, is, when properly regimented, revealed as embodying 'a non-sequitur of numbing grossness'."

(iii) We do not know that for all A, either it is possible to prove A or it is possible to prove the negation of A. (This is a principle of epistemic modesty.)

We have just briefly recalled the arguments in favor of the Knowability Principle. The principle of epistemic modesty is reasonable as well: there are large classes of sentences for which we do not possess any decision procedure, where by a decision procedure, we mean an effective method yielding a proof of A if A holds and a proof of the negation of A, if the negation of A holds. It follows then that we should reject the first claim, namely that we are entitled to assert LEM in full generality.

On the contrary, it is clear that LEM holds for those classes of sentences for which we do possess a method for deciding them.[5] But the argument above shows that it is not sound in general to assume LEM, and that we should hold on to it *only when we are concerned with decidable classes of sentences*. The anti-realist is therefore a logical revisionist in so far as she draws a line between those statements for which LEM can be asserted and those for which it cannot. And decidability is the criterion used to draw this line, because decidability is both necessary and sufficient for us to be entitled to assert LEM.

Without entering into the details of Salerno's argument, we shall be content with this presentation of high-level revisionism. Let us consider now what we have called *low-level revisionism*. There is a normative component that any theory of meaning based on assertibility conditions should abide by: What can be inferred from a given sentence should not go beyond what is required in order to be entitled to assert it. This principle of *harmony* takes a precise form in the setting of natural deduction which is used to provide the meaning of logical and mathematical expressions.[6] In natural deduction, the assertability conditions—the conditions for being in position of asserting a statement—are given by the introduction rules, and the corresponding "exploitability conditions"—what can be inferred from a statement—are given by the elimination rules. The principle of harmony has it that every detour consisting in an introduction rule followed by an elimination rule for the same expression should be eliminable.

It turns out that the rules of intuitionistic logic satisfy harmony. But, under the assumption that the calculus should not allow for multiple conclusions, LEM or other principles yielding classical logic such as double negation elimination cannot be added in such a way that harmony obtains. We should therefore reject classical logic in favor of intuitionistic logic, because the latter, but not the former, is

[5] A class of sentences is decidable if for any sentence in the class, a speaker is always in position to know whether she can assert it or not.

[6] Natural deduction can come either as a mono-conclusion system or as a multiple conclusions system. Harmony rules out classical logic only if the system admits only of a single conclusion at a time. Dummett has argued that using multiple conclusions is not ok because this presupposes an (unsound) classical understanding of disjunction. This point has been recently challenged by Restall (see [16]). Restall proposes conceptual foundations for a system with multiple conclusions based on two primitives, assertion and denial. This seems to us to be a very promising response to the anti-realist challenge againt classical logic, though a discussion of Restall's arguments would lead us beyond the scope of this paper.

satisfactory from a normative perspective. Low-level revisionism is thus based on a proof-theoretic semantics. It is important to note that this is a two-stage path. First, one endorses an assertibility-conditions theory of meaning. Then, as a by-product, classical logic is disqualified and intuitionistic logic is justified.

A striking feature of logical revisionism along the lines of moderate anti-realism emerges when one compares high-level revisionism with low-level revisionism: both lead to the very same conclusion, namely that LEM should be rejected. On the one hand, High-level revisionism discards the principle of bivalence, leaving us with no reason to accept LEM. On the other hand, low-level revisionism justifies intuitionistic logic, which may be construed as classical logic minus LEM. Not only low-level revisionism is consistent with high-level revisionism, but it does not advocate any further departure from classical logic than the one which is required by high-level revisionism. There is thus some kind of "meta-harmony" between the two levels of revisionism. As Dummett puts it [8, p. 75], "A theory of meaning in terms of verification is bound to yield a notion of truth for which bivalence fails to hold for many sentences which we are unreflectively disposed to interpret in a realistic manner". Low-level revisionism shows that a theory of meaning in terms of verification does yield the logic it is bound to yield on account of high-level revisionism.

3.2.2 Radical Anti-realism

The radical anti-realist shares the basic tenets of the anti-realist, but she thinks her colleague is too shy when it comes to putting epistemic constraints on truth. As a consequence, the radical anti-realist will be a revisionist too, but she will be an even more radical one. We will start by explaining how and why the basic principles of anti-realism are radicalized, and then we will examine the consequences of this move for logic.

3.2.2.1 Decidability in Principle and Decidability in Practice

According to the anti-realist, if a statement is true, one has to be able to recognize that it is true. And for LEM to hold, statements have to be decidable. But what does it mean to say that one has to be able to recognize that something is true, or to decide if a statement is provable or disprovable? On the one hand, the cognitive abilities of a not that gifted sophomore are certainly not the absolute norm by which truth should be constrained. On the other hand, the limitless powers of the divine intellect are not a reasonable candidate either. If truth is only constrained by what God can do, and if God can do anything, this is a cheap constraint indeed.

What are the norms by which recognizability of truth are to measured? The moderate anti-realism does not choose God's point of view; indeed Dummett acknowledges that if these norms were taken to be those of God, realism and anti-realism would conflate into one and the same position. However, moderate anti-realism is still quite *liberal* with respect to these epistemic constraints: for a set of sentences to be a decidable class, it is only required that such sentences might be decidable *in principle* by a creature with a finite mind, that is by finitary mechanical procedures.

Now, the problem is that moderate anti-realism has to face some kind of revenge. If truth has to be epistemically constrained in order to satisfy manifestability requirements, these constraints have to be strong enough to guarantee that knowledge of truth is manifestable. But think of a set of sentences which is decidable in principle, but such that the truth or falsity of some sentences can only be established by methods which are *practically* out of reach. In that case, what is there to be exhibited? If the decision procedure cannot actually be used and applied, in which sense would knowledge of these methods be any more human than God's knowledge? What would it mean to manifest such a knowledge, or to be able to acquire it? Thus it seems that for such a set of sentences the moderate anti-realist fails to satisfy the requirements that she has herself advertised against realism. Granting this point, moderate anti-realism appears as an *unstable* position: epistemic constraints on truth might be discarded right at the beginning, but if there are such constraints, they should be taken seriously and they should be measured by decidability *in practice* instead of decidability *in principle*.

3.2.2.2 The Radical Anti-realist Crush on Substructural Logic

There have been various attempts to implement this radicalization, among which strict finitism is one of the most famous (as elaborated for example in [20]). In this paper however, we shall focus on another version of radical anti-realism recently advocated in [5].[7] According to Dubucs and Marion, the outcome of the radical anti-realist revision procedure should no longer be intuitionistic logic. They claim that substructural logics (see [15] for a thorough introduction), and more precisely linear logic, are more faithful to the basic insights of anti-realism than intuitionistic logic.

Let us see why. In standard presentations of sequent calculus, different types of rules are distinguished. There are on the one hand the *logical rules*, which make for the introduction of logical connectives, and there are on the other hand the *structural rules*, like the rules of Weakening and Contraction, which correspond to properties of the consequence relation itself.[8] Here are Weakening and Contraction:

---

[7] We shall not explain here in any detail why we favor this approach over strict finitism. Basically, we agree with the arguments by Dubucs and Marion against strict finitism. Specifying by brute force what it means to be feasible—say it means "being doable in less than n steps of computation", or "doable in a reasonably small number of steps"—is bound to lead to soritic paradoxes. Despite the criticisms that we develop on here, we take the proposal by Dubucs and Marion to be the most attractive one among various versions of radical anti-realism, precisely because it aims at getting a non-stipulatory grip on feasibility.

[8] Note that this distinction is not tied to the adoption of sequent calculus as a proof system. A similar point could be made using, say, natural deduction, tableaux methods or a dialogical setting. Arguably, any good framework for proofs is able to distinguish between abstract properties of the consequence relation, that may or may not be used in proofs, and the mere characterization of logical connectives by logical rules.

$$\frac{\Gamma \vdash \Delta}{A, \Gamma \vdash \Delta}\ \textit{Weakening (left)} \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, A}\ \textit{Weakening (right)}$$

$$\frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta}\ \textit{Contraction (left)} \qquad \frac{\Gamma \vdash \Delta, A, A}{\Gamma \vdash \Delta, A}\ \textit{Contraction (right)}$$

The radical anti-realist's idea is that some substructural rules contain crucial elements of epistemic idealization. Hence, in order to "unidealize" logic from an epistemic point of view, one should control these structural rules. A new logical revisionism follows: the claim is now that a substructural logic like linear logic is justified from an anti-realistic point of view.

This is the radical anti-realist crush on substructural logic, and the aim of this paper is to evaluate it. One may basically evaluate such a proposal from two points of view: from the point of view of the *principles* that could lead to it and from the point of the *consequences* that would result from its endorsement. We will proceed to the evaluation from both points of view and deal with the following two questions:

- How can one live without structural rules?
- Why should one divorce from them?

3.3 Life Without Structural Rules

Opponents to semantic anti-realism have always been prompt to notice that there is something paradoxical in the anti-realist's position. The anti-realist bases her rejection of realism on slogans such as "meaning is use" and she ends up with a proposal to revise usage. Stated in polemical terms, this amounts to saying that the anti-realist's attitude towards use is opportunist. She invokes use when it is useful to do so and repudiates it when needed. Anti-realists grant the existence of such a tension, but claim that there is nothing preposterous in it. However, M. Dummett admits that the greater the revisions, the less plausible the theory, because "the principal purpose of a theory of meaning is to explain existing practice rather than to criticize it."[9]

Obviously, this tension will be all the more vivid in the context of radical anti-realism. If its advocate grants with Dummett that an increase in departure from "existing linguistic practice" yields a decrease in the theory's plausibility, then she cannot but hope that the shift to substructural logic is not a too dramatic revision.

How should we assess the acceptability of a revisionist proposal with respect to standard use? In the case of logical connectives, we take it that the most elementary inferences that speakers accept as part of a characterization of what these connectives mean should be recognized as valid. Of course what these inferences are is a matter of debate, but we shall argue that, on any account of what are the basic

---

[9] [8, p. 75], "What is a Theory of Meaning? (II)".

meaning-constitutive inferences for the logical connectives, the revision in point is quite severe. Our concern is related to a well-known feature of substructural logics, namely the so-called phenomenon of splitting of logical connectives. Let us consider two pairs of rules for conjunction in (intuitionistic) sequent calculus:

Introduction on the left	Introduction on the right	
$\dfrac{\Gamma, A \vdash C}{\Gamma, A \wedge B \vdash C}$ $\quad$ $\dfrac{\Gamma, B \vdash C}{\Gamma, A \wedge B \vdash C}$	$\dfrac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B}$	additive $\wedge$
$\dfrac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C}$	$\dfrac{\Gamma_1 \vdash A \quad \Gamma_2 \vdash B}{\Gamma_1, \Gamma_2 \vdash A \wedge B}$	multiplicative $\wedge$

One can check that the two pairs of rules are equivalent, in the sense that each one can be derived from the other. But this derivation resorts crucially to the structural rules of Weakening and Contraction. Without such rules, the equivalence does not hold. Therefore, in a context in which the structural rules are not valid, one gets two different conjunctive connectives: one that corresponds to the first pair of rules, and the other that corresponds to the second pair of rules. The latter is called *fusion* in the relevant logic literature, *multiplicative conjunction* or *times* (notation: $\otimes$) in the linear logic literature. The former is called *additive conjunction* or *with* (notation: &) in the linear logic literature.

The two right-introduction rules make the difference between the two conjunctions salient. In the additive case, there is one antecedent Γ which is common to $\Gamma \vdash A$ and $\Gamma \vdash B$, whereas in the multiplicative case, the antecedents may be different. For our discussion, the main question is to know what are the connections between these two connectives and our pre-theoretical notion of conjunction. Let us consider the two connectives in turn:

(i) $\otimes$: one can easily show that the following sequent is derivable:

$$A, B \vdash A \otimes B$$

which seems to be a highly desirable feature for a conjunction: to get A and B, I just need both A and B. Furthermore, the interaction between $\otimes$ and $\rightarrow$ satisfies the so-called residuation property:

$$A \rightarrow (B \rightarrow C) \equiv (A \otimes B) \rightarrow C$$

If A implies that B implies C, then I can get C from A and B, and vice versa. This is quite reasonable. But note that sequents of the form

$$A \otimes B \vdash A$$

are *not* derivable. From a pre-theoretical point of view, this behavior of $\otimes$ is weird. If I can show that A and B, why should not I be able to assert A?

(ii) &: the additive conjunction has welcome features as well. In particular, the following sequent is derivable:

$$A \& B \vdash A$$

But it is no longer possible to derive:

$$A, B \vdash A \& B$$

To put it bluntly, multiplicative conjunction seems to describe nicely the conditions under which a conjunction can be asserted but not what can be inferred from a conjunction. On the contrary, additive conjunction seems to describe nicely what can be inferred from a conjunction but not the conditions under which a conjunction can be asserted.

Several replies are available to the radical anti-realist. She can argue that one of the two connectives is the true one. However, given what we have just said, this does not seem very plausible. Another reply would be to bite the bullet and consider that linear logic refines on our pre-theoretic use of conjunction which is ambiguous. From this point of view, contrary to what the layman thinks, there is no single well-defined notion of conjunction. The layman might be wrong, in the sense that our best theory of meaning might have among its consequences that "and" is indeed ambiguous. However, note that "and" fails the standard linguistic test for ambiguity, namely cross-linguistic disambiguation. "Bank" in English is ambiguous between some place where I can get money and some slope beside a river where I can sit and wait for a fish to bite the hook. One good reason to think that there are two different lexical entries for "bank" is that in other languages, like French, there are two different words for that, namely "banque" and "rive". We do not know of any spoken language in which there would be two different words for "and", one corresponding to additive conjunction, the other one to multiplicative conjunction.

Whatever the reply the radical anti-realist chooses, we take this to show that life without structural rules is not easy. Arguably, this does not constitute a knockdown argument to reject the radical anti-realist's proposal. But considering these difficulties, she better have very good reasons to divorce from structural rules. In other words, the reasons for rejecting structural rules have to be pretty strong in order to balance the cost of living without them. Hence, we now turn to the assessment of these reasons.

3.4 The Anti-realist Justification of Substructural Logic

Let us scrutinize more closely reasons given by radical anti-realists for dropping some of the structural rules. The way we proceed will follow our reconstruction of moderate anti-realism. We will first consider high-level revisionism, and then low-level revisionism.

3.4.1 High-Level Revisionism

Radical anti-realism ensues from a strengthening of the epistemic constraints. What does high-level revisionism amount to in this context? A striking feature of the version of radical anti-realism we are discussing is that its denial of moderate anti-realism's idealizations leads to the rejection of *new* logical laws (the structural rules, instead of just LEM). Something here is puzzling. Requiring decidability in practice makes the class of problematic sentences larger, but why should such a shift have revisionist implications of a *different* kind?

Our point is the following. As we have stressed in the first section, the disagreement between realists and moderate anti-realists concerns classes of undecidable sentences: the moderate anti-realist rejects the disputed logical principle, namely LEM, precisely for those classes. Let us assume that there is an argument Π which relies on the principle that truth should be epistemically constrained and which does show that, for undecidable classes, LEM does not hold (Π is the kind of argument that we have mentioned in Section 3.2). Let us assume furthermore that, in the previous principle, decidability in principle should be replaced by decidability in practice. As a consequence, Π is likely to be turned into a stronger argument Π', which shows that, for domains which are undecidable in practice, LEM does not hold. What is crucial here is that the shift from Π to Π' does not change the logical law (i.e. LEM) that is under dispute, but changes the scope of the domain of validity of that law. Arguably, the domain of validity of LEM becomes more restricted: the law is no longer valid for every domain[10] which is decidable in principle, but only for domains which are decidable in practice. Domains which are decidable in principle but not in practice will fall outside of the scope of the law.

The point has actually been made by C. Wright in his book on strict finitism:

> ... whereas the intuitionist is content to regard as determinately true or false any arithmetical statement whose truth value can be effectively computed, at least "in principle", the strict finitist will insist that the principle of Bivalence is acceptable only for statement the verification or falsification of which can be guaranteed to be humanly feasible. (in [20, p. 108])

As a consequence, the landscape to be drawn should not be:

realism	$\Rightarrow$	classical logic
moderate anti-realism	$\Rightarrow$	intuitionnistic logic
radical anti-realism	$\Rightarrow$	linear logic

[10] By a decidable domain, we mean a domain such that for any tuples of objects in the domain, for any predicate, a speaker is always in position to know whether the predicate applies to the objects. This is in keeping with our previous use of decidable as a property of classes of sentences.

but rather:

realism	$\Rightarrow$ LEM for all domains
moderate anti-realism	$\Rightarrow$ LEM restricted to domains decidable in principle
radical anti-realism	$\Rightarrow$ LEM restricted to domains decidable in practice

Restricting LEM to decidable domains and choosing intuitionistic logic is perfectly coherent. Since LEM is not valid in full generality, one should choose a logic such as intuitionistic in which the principle is not a theorem. It just happens that for some special domains, the decidable ones, LEM can be used, because of the property these domains have. The same is not true with restricting LEM to a subclass of decidable domains and choosing linear logic: the shift from intuitionistic logic to linear logic cannot be analyzed as a consequence of further restricting the validity of LEM.

To sum up, it is clear that, from a high-level perspective, radical anti-realism is bound to yield an even more radical revision. Nonetheless, it is not yet clear why the nature of this revision should be any different from the one advocated by the moderate anti-realist. Therefore, if the radical anti-realist is to propose a new kind of logic, such as linear logic, this justification has to take place from a low-level perspective. And, in any case, the convergence between high-level and low-level revisionism, which was a nice feature of moderate anti-realism, will be lost.

3.4.2 Low-Level Revisionism

Now we shall turn to low-level revisionism. The question is: can radical anti-realism do for linear logic what moderate anti-realism does for intuitionist logic? That is, can moderate anti-realism both vindicate the rules of a substructural logic and provide reasons to reject stronger systems?

To answer this question, some preliminary remarks are in order. First, moderate anti-realists do put forward a criterion, the criterion of harmony, which discriminates between acceptable and unacceptable pairs of rules. Radical anti-realists have to provide an analogous but more demanding criterion. Second, the radical anti-realist and the moderate anti-realist do not seem at first sight to talk about the same thing. The moderate anti-realist focuses on logical rules *stricto sensu*,[11] whereas the radical anti-realist targets structural rules. Harmony is tailor-made for logical rules in a natural deduction format.

Thus in order to provide a complete justification, the moderate and the radical anti-realist have to propose admissibility criteria both for structural and for logical

[11] To our knowledge, Dummett does not discuss the validity of structural rules at all. One contingent reason might be that he uses systems of natural deduction in which structural rules are built-in rather than introduced as genuine rules.

rules. Our aim will be to sketch the ways in which these expectations could be fulfilled. Four criteria are needed, as can be seen in the following tabular:

	moderate anti-realism	radical anti-realism
logical criterion	harmony	?
structural criterion	?	?

Let us consider first the admissibility criteria for logical rules (upper line). Right now, there is only one cell whose content is obvious. The moderate anti-realist takes the principle of harmony as a requirement on logical rules. It is clear that whatever is required for the moderate anti-realist is also required for the radical anti-realist. So the radical anti-realist's logical admissibility criterion has to be at least as strong as the principle of harmony.

But there is a question concerning the means which are available in order to eliminate the detours. For the moderate anti-realist, structural rules are available. But the radical anti-realist rejects the structural rules: therefore, when she requires harmony, she will also require that detours can be eliminated *without resorting to structural rules.*

For example, consider the following rules for conjunction:

$$\frac{\Gamma \vdash A \quad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash A \wedge B} \wedge - intro \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \wedge - elim \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} \wedge - elim$$

These rules are harmonious as far as the moderate anti-realist is concerned. Consider a detour consisting in an introduction rule followed by an elimination rule on the same conjunction:

$$\cfrac{\cfrac{\Gamma \vdash A \quad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash A \wedge B} \wedge - intro}{\Gamma, \Gamma' \vdash A} \wedge - elim$$

The detour can be eliminated and replaced by one application of Weakening:

$$\frac{\Gamma \vdash A}{\Gamma, \Gamma' \vdash A} Weakening\ (left)$$

Either the anti-realist is able to eliminate all such detours without resorting to Weakening,[12] or she has to choose so-called 'pure' pairs of introduction and elimination rules for which the detours can be eliminated without structural rules. This is the case of the rules for additive conjunction or multiplicative conjunction. But mixing the two would not work.

[12] It is actually hard to see how this could be done.

To sum up, the most natural logical criterion for the radical anti-realist is nothing but a strengthened version of harmony, in which the use of structural rules is banned. The following picture arises:

	moderate anti-realism	radical anti-realism
logical criterion	harmony	strong harmony
structural criterion	?	?

Let us consider normative criteria for structural rules. What could the moderate anti-realist say? To start with, no analogue of the principle of harmony is at hand. Roughly, harmony is meant to show that "nothing new" is introduced, in so far as harmony implies, at least in appropriate contexts, conservativity. But structural rules *do* introduce some new proof means: there are things which can be proved with structural rules which cannot be proved without them. For example, if we consider a given atomic basis B in the sense of Prawitz (i.e. a set of mono-conclusion sequents containing only atomic sentences), it is in general possible that there is a sequent S which is not in B and which can be proven from B by using Weakening or other structural rules.

This means that one has to provide a full-fledged justification of structural rules, which does not rely on some sort of eliminability arguments. We suggest the following principle:

Preservation of Effectivity

A structural rule of the form

$$\frac{\Gamma \vdash A}{\Gamma' \vdash A}$$

is admissible iff,

if there exists an effective means to transform justifications for all sentences of Γ into a justification for A, then there exists an effective means to transform justifications for all sentences of Γ' into a justification for A

The principle of preservation of effectivity is in the spirit of the BHK interpretation of logical constants. It is applied here at the meta-level to the consequence relation represented by the turnstile. Because of the close connection between the consequence relation in the meta-language and implication in the object-language, it comes as no surprise that our principle mirrors the BHK clause for implication. The anti-realist demands that proofs provide us with effective justifications, nothing less, but nothing more. Therefore, the principle of preservation of effectivity seems to express both necessary and sufficient conditions for the admissibility of structural rules.

This principle validates the standard structural rules. Weakening is admissible:

$$\frac{\Gamma \vdash A}{\Gamma, B \vdash A}$$

If one has an effective method to get a justification for A from justifications for sentences in Γ, one has also an effective method to get a justification for A from these justifications plus a justification for B. One just has to *discard* the unnecessary justification for B.

Contraction is admissible as well:

$$\frac{\Gamma, A, A \vdash A}{\Gamma, A \vdash A}$$

If one has an effective method to get a justification for A from justifications for sentences in Γ plus a justification for A and another one for the same sentence A, one has also an effective method to get a justification for A from the justifications for sentences in Γ and the remaining justification for A. One just has to *duplicate* the remaining justification for A whenever needed. (It is crucial here that the effective method provided for the upper sequent has to work whatever justifications for A are given.)

Exchange is admissible as well:

$$\frac{\Gamma, A, B, \Gamma' \vdash A}{\Gamma, B, A, \Gamma' \vdash A}$$

If one has an effective method to get a justification for A from justifications for sentences in Γ, for A, for B and for sentences in Γ', one has also an effective method to get a justification for A from justifications for sentences in Γ, for B, for A and for sentences in Γ'. One just has to look for the required justifications in the right place: the order on the left hand side of the sequent does not matter.

By contrast, let us have a quick look at the following rule, which we might call $stronk$:

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash B}\ \textit{stronk}$$

stronk is some kind of strengthening which exhibits the same misbehavior as *tonk*. Let us assume that we have an effective procedure to get a justification for B from justifications for sentences in Γ and a justification for A. It might be the case that such a procedure makes an essential use of the justification provided for A, and that there is no effective procedure giving us a justification for B on the basis of Γ alone. Hence, *stronk* is not an admissible rule.

Preservation of effectivity vindicates the conservative attitude of moderate anti-realism towards structural rules. Our point in introducing this principle has only been to show that the gap left by the overlooking of structural rules could be bridged, and to prepare the ground for the discussion to come concerning radical anti-realism. We thus get the following picture:

	moderate anti-realism	radical anti-realism
logical criterion	harmony	strong harmony
structural criterion	preservation of effectivity	?

Let us turn now to the crucial last part of the discussion, namely the radical anti-realist view of structural rules. More precisely, we will discuss in turn three ways of filling the last blank in our tabular:

(i) Token preservation
(ii) Preservation of local feasibility
(iii) Preservation of global feasibility

(i) As noted above, the crucial claim of the radical anti-realist is that one should reject *both* the Weakening rule and the Contraction rule. By contrast, the radical anti-realist has no quarrel with the Exchange rule.[13] We will start by proposing a criterion which is designed to account for exactly this attitude towards structural rules, and then see whether it can be justified from the anti-realist perspective. Here is the suggestion:

Principle of Token Preservation:

A structural rule of the form

$$\frac{\Gamma \vdash A}{\Gamma' \vdash A}$$

is admissible if, for every formula B, the number of tokens of B is the same in Γ and in Γ'.

It is easy to see that the principle of Token Preservation rules out the Weakening and Contraction rules. On the contrary, the Exchange rule is justified according to it. Of course, pathological *stronk* is ruled out as well.

The Principle of Token Preservation thus seems to mirror adequately the radical anti-realist attitude towards the different structural rules. Of course, this is not enough: adopting the principle just on the ground that it yields the desired result would be an entirely ad hoc move, if there were no justification of it on account of the basic tenets of radical anti-realism.

Here is an attempt at such a justification of the principle. Linear logicians sometimes motivate their logic by providing an informal semantics for their calculus in

[13] One might ask why this is so. What about a super-radical anti-realist who would dismiss Exchange as well as Weakening and Contraction? Logical systems exist which would fulfill the super-radical anti-realist dreams. Non-commutative linear logic is one of them. The super-radical anti-realist would have to argue that the order in which evidence is given is not neutral with respect to our ability to draw conclusions on the basis of that evidence. A super-radical stance on feasibility might support this view. Think of an agent being asked whether it is valid to infer A from $A, A_1, \ldots, A_{100000}$, as compared to being asked whether it is valid to infer A from $A_1, \ldots, A_{52227}, A, A_{52228}, \ldots, A_{100000}$, the protocol being such that all A_is for $1 \leq i \leq 100000$ are actually presented to the agent. It might be feasible to answer the first question but not the second.

terms of *resources and resource consumption* (see [10]). On this interpretation, types of formulas stand for types of resources and a sequent of the form $\Gamma \vdash A$ expresses the fact that one can get an object of type A from resources corresponding to the elements of Γ. In this perspective, the Contraction rule becomes problematic because it says, for instance, that if one may get something of type A from two resources of type B, then one may get something of type A with just one resource of type B. But being able to buy a pack of Marlboros with two bucks does not guarantee that one can buy a pack of Marlboros with just one buck. It seems to us that this "resource interpretation" is conceptually defective from an anti-realist perspective. The reason is the following one. The resource interpretation is based on a sort of causal reading of the turnstile, formulas standing for types of objects such that the consumption of some can result in the production of others. In terms of the previous example, my two bucks can be traded for a pack of cigarettes. But the radical anti-realist is concerned with epistemic constraints on speakers. It is fallacious to assimilate the two perspectives. Of course, inference steps have a cognitive cost, and it might well be the case that some inference steps have a cognitive cost significantly higher than some others, so that a radical anti-realist should be particularly reluctant to admit them in her favorite logic. But, nonetheless, cognitive resources are not on a par with consumption goods. A justification does not disappear when I use it to build another justification in the same way that buying a pack of cigarettes makes a dollar or two disappear out of my pocket. For this reason, we do not think that a rejection of structural rules can be based on the "resource interpretation" of linear logic.

(ii) As a consequence, the radical anti-realist has to look for another kind of structural criterion of admissibility. The most promising line of thought consists in radicalizing what we have called above Preservation of Effectivity. It is very natural to do so since the basic criticism that the radical anti-realist addresses to her moderate cousin is that one should require not only effectivity in principle but effectivity in practice or feasibility. Therefore, one might wish to consider the following twist on Preservation of Effectivity:

Preservation of Local Feasibility

A structural rule of the form

$$\frac{\Gamma \vdash A}{\Gamma' \vdash A}$$

is admissible if, if there exists a feasible means to transform justifications for all sentences of Γ into a justification for A, there exists a feasible means to transform justifications for all sentences of Γ' into a justification for A

If one accepts to consider this requirement as a consequence of the radical anti-realist's position, the crucial question is: does Preservation of Local Feasibility gives us a reason to reject the Weakening and Contraction rules?

Consider Weakening. Let us assume that we have an effective and *feasible* method to get a justification for A from justifications for sentences in Γ. What kind of method do we get for extracting a justification for A from these justifications plus a justification for B? The same as before, except that one has now to discard the unnecessary justification for B. Why on earth would this not be feasible? After all, the idea is just to do as if no new justification would have been given, and to stick to the good old feasible procedure, *even if* we could now try to use in non-feasible ways the new justification we have just been provided with. But note that for Preservation of Local Feasibility to hold, it is sufficient that the property of being in possession of a feasible procedure is preserved. It would not make sense to require that all possible procedures are feasible, since it is clear that there might always be non-feasible ways of doing feasible things (try to unload a ton of sand from a truck with a pitchfork instead of a spade).

The same goes with Contraction. Let us assume that we have an effective method to get a justification for A from justifications for sentences in Γ plus a justification for B and another one for the same sentence B. As we said, to get a transformation procedure corresponding to the lower sequent, one has to be able to "re-use" the justification provided for B. Again, why on earth would this not be feasible? By assumption, the justification for B has to be "simple" enough to be dealt with in the transformation procedure. But if this is the case, why would it suddenly cease to be "simple" enough to be re-used?

(iii) The radical anti-realist might reply that it is intuitively clear that it is harder to get a proof of A from Γ and B than from Γ alone, because one has to take into consideration that B might be necessary to prove A from Γ. If we do not grant this point, the anti-realist might claim that our disagreement reflects a brute conflict of intuitions and that our arguments do not bear upon his analysis of what the logic of feasible proofs is. But by saying so, the radical anti-realist makes a significant shift from the question of the preservation of feasibility for transformation procedures to the question of the feasibility of establishing that, let's say, A follows from Γ. The latter question concerns the complexity of the consequence relation, i.e. a global property of the logic. Instead of discussing whether some property is locally preserved by applying Weakening or Contraction, we are now dealing with a different idea, namely that admissibility of structural rules should be judged on the basis of their effects on logical systems. This global requirement of feasibility bears upon the calculus as a whole, as opposed to the local requirement we had introduced. The new principle could be spelt out in the following way:

Preservation of Global Feasibility

A set of structural rules S preserves global feasibility w.r.t. a set of logical rules L, iff, if $\vdash_L$ is feasible, then $\vdash_{L+S}$ is feasible as well.

For the radical anti-realist's intuitions to be mathematically vindicated, the complexity of the consequence relation of a logic without structural rules should be lower than the complexity of the consequence relation of the same logic to which structural rules have been added. In particular, the radical anti-realist seems to be committed to the claim that it is feasible to establish whether A follows from Γ in linear logic whereas it is not the case in intuitionistic logic.

However, at this point, some well-known results in computability theory and complexity theory preclude such vindication. As a matter of fact, the consequence relation is decidable and PSPACE-complete in intuitionistic logic[14] but is undecidable in full linear logic. If one drops the exponentials, linear logic becomes decidable but is still PSPACE-complete. Furthermore, if one shifts from full linear logic to affine logic, i.e. linear logic plus the Weakening rule, one goes from undecidability to decidability (see [11]). As a consequence, there seems be no correlation between the feasibility of establishing that A is a consequence of Γ and the rejection of structural rules. To the contrary, such a rejection sometimes makes matter worse (see also [17], who makes a similar point against strict anti-realism).

3.5 A Way Out for Radical Anti-realism?

Up to now, our analysis of radical anti-realism has led to two main claims:

- Radical anti-realists have to provide a low-level justification of their choice of linear logic.
- Such a justification is still to be provided. In particular, three direct attempts have been examined and shown to fail.

Our criticism of the three putative requirements on structural rules (token preservation, preservation of local feasibility and preservation of global feasibility) are not on a par. For purely mathematical reasons, the idea of preserving global feasibility seems to us to be misguided. The problem with token preservation is its lack of conceptual support: the anti-realist does not explain why justifications should share the properties of consumption goods. In a sense, preservation of local feasibility can be construed as some kind of conceptual support in favor of token preservation. However, our criticism of local feasibility suggests that, on this account, an informal notion of justification is not likely to invalidate structural rules.

To be fair, the radical anti-realist could blame our failure to see why structural rules are problematic on our informal analysis of justifications. She could claim that, on her view of what justifications are, two justifications can be (substantially) better than one. Now, of course, such a view has to be spelled out. Some help could come from the proof semantics that have been given for linear logic: as long as such a semantics could be considered to provide a formal counterpart for a reasonable notion of justification, it would provide intuitive counterexamples to the admissibility of structural rules and to the provability of the sequents that can be derived by using them.

The game semantics which have first been proposed by Blass [1] provide a case in point. In this setting, justifications are defined as winning strategies for a designated player on two-players games (the two players are P for Player and O for opponent, where P is the designated player who tries to "verify" the formula). Given games for atomic formulas, each complex formula is associated with a mathematically

[14] We hereby mean that the problem of deciding whether a pair of formulas stand in this relation or not is PSPACE-complete.

well-defined game, which is defined by recursion on the syntax. Provability of a sequent $\vdash A$ amounts to the existence of a winning strategy for P no matter what the atomic games are. Blass gives an example of an infinite game for an atom G such that O has a winning strategy for $G \otimes G$ even though she does not have one for G alone. Intuitively, this accounts for the fact that "G and G" (where "and" is multiplicative conjunction) can be harder for me to justify than "G" alone (my opponent might be able to refute my claim that G and G though she is not able to refute my claim that G).

However, a crucial feature of Blass' semantics is the use of infinite two-players game, which are responsible for the failure of determinacy and hence for the differences between G and $G \otimes G$. From an anti-realist perspective, the meaning of infinite "justificatory debates" is not clear. But there are other ways to lose determinacy. In particular, a natural anti-realist constraint on strategies would consist in feasibility requirements: a justification for a formula G should not be any kind of winning strategy, but a feasible one, where feasibility would be captured in terms of a measure of complexity on strategies (say, we should only consider strategies computable by finite automata of a given size to reflect the cognitive limitations of the agent, see [13] for more on this). In a given finite game, one of the two players has some winning strategies, but all these strategies may well fall outside of the class of the feasible ones. Hence the failure of determinacy.

To put it bluntly, there is on the one hand a story told by the radical anti-realist explaining to us why we should worry about intuitionistic logic and its anti-realist foundations. On the other hand, there are various available semantics for linear logic which show for what kind of notions of justifications structural rules can fail to be admissible. A thorough vindication of linear logic by radical anti-realism would have to make this story and one of these semantics meet. Our point in the previous section was to suggest that this is by no way an easy task, and that an elaborate notion of justification is needed. Our suggestion in the present section is that game semantics together with complexity constraints on available strategies could be a reasonable candidate.[15]

3.6 Conclusion

As the name says, radical anti-realism is a radicalized version of anti-realism. In terms of conceptual motivations, it rests upon the idea that anti-realism stops halfway in its request for feasibility. Radical anti-realism demands feasibility in

[15] As a witness of the recent interest of strict anti-realists in games, see [12]. In discussion, Greg Restall has suggested another possible line of defense for strict anti-realism. By the well-known Curry-Howard correspondence between proofs and programs, proofs in intuitionistic logic can be turned into computable functions represented by λ-terms. Dropping structural rules shrinks the class of typable λ-terms (Weakening allows for empty binding and Contraction for multiple binding). The strict anti-realist would have a point if she could show that the class of functions typable in linear logic corresponds to a more feasible class of functions than those typable in intuitionistic logic.

practice instead of mere feasibility in principle. In terms of logical revisionism, radical anti-realism advocates a more radical departure from classical logic. Its favored logical systems are substructural logics, and in particular linear logic. We have proposed a systematic discussion of the arguments from conceptual motivations to logical revisionism, comparing the case for the move from feasibility in practice to substructural logics to the case for the move from feasibility in principle to intuitionistic logic.

Only one of the two main arguments available to moderate anti-realists is available to radical anti-realists. Direct objections to the law of excluded middle will not speak in favor of systems strictly weaker than intuitionistic logic. The other option is to lay down explicit admissibility criteria for logical and structural rules. Harmony is the famous criterion put forward by Dummett for logical rules. There is no such standard proof-theoretic criterion for structural rules. To make up for this, we have suggested three possible criteria that might appeal to radical anti-realists.

These criteria were based on a straightforward implementation of the idea of feasibility in practice in a broadly "BHK-like" framework. None of them seems to us to be satisfactory, so that one lesson of our efforts is that the move from radical anti-realism to substructural logics is probably more problematic than the corresponding move from anti-realism to intuitionistic logic. In the very last section of this paper, we suggest that a promising alternative would be to construe proof-theoretic semantics,[16] such as game semantics, as implicitly providing a logical analysis of justifications on which feasibility requirements could be imposed. The next step, which falls outside the scope of this paper, obviously is to work out the details of such a proposal.

Acknowledgements We thank the audience of the colloquium *(Anti-)Realisms: Logic & Metaphysics* held in Nancy (July 2006), and especially J. Dubucs, M. Marion and G. Restall. We also wish to thank an anonymous referee for further thoughtful comments. The first author acknowledges support from ESF (the Eurocores LogICCC program).

References

1. Blass, A. 1992. "A Game Semantics for Linear Logic." *Annals of Pure and Applied Logic* 56:183–220.
2. Cogburn, J. 2002. "Logical Revision Re-Revisited: On the Wright/Salerno Case for Intuitionism." *Philosophical Studies* 110:231–48.
3. Cogburn, J. 2003. "Manifest Invalidity: Neil Tennant's New Argument for Intuitionism." *Synthese* 134:353–62.
4. Dubucs, J. 2002. "Feasibility in Logic." *Synthese* 132:213–37.

[16] Philosophers sometimes mean by proof-theoretic semantics a semantics for mathematical sentences in terms of conditions of provability. Proof-theorists mean by proof-theoretic semantics semantic accounts of the nature of proofs (including criteria of identity for proofs).

5. Dubucs, J., and M. Marion. 2003. "Radical Anti-realism and Substructural Logics." In *Philosophical Dimensions of Logic and Science*, edited by A. Rojszczak and J. Cachro, 235–49. Dordrecht: Kluwer.
6. Dummett, M. 1978. *Truth and Other Enigmas*. London: Duckworth.
7. Dummett, M. 1991. *The Logical Basis of Metaphysics*. London: Duckworth.
8. Dummett, M. 1993. *The Seas of Language.* Oxford: Clarendon Press.
9. Girard, J.-Y. 1987. "Linear Logic." *Theoretical Computer Science* 50:1–102.
10. Girard, J.-Y. 1995. "Linear Logic: Its Syntax and Semantics." In *Advances in Linear Logic*, edited by J.-Y. Girard, Y. Lafont, and L. Regnier, 1–42. Cambridge, MA: Cambridge University Press.
11. Lincoln, P. 1995. "Deciding Provability of Linear Logic." In *Advances in Linear Logic*, edited by J.-Y. Girard, Y. Lafont, and L. Regnier, 109–22. Cambridge, MA: Cambridge University Press.
12. Marion, M. 2005. "Why Play Logical Games?" In *Games: Unifying Logic, Language, and Philosophy*, edited by O. Majer, A.-V. Pietarinen, and T. Tulenheimo, 3–26. Dordrecht: Springer.
13. Neyman, A. 1998. "Finitely Repeated Games with Finite Automata." *Mathematics of Operation Research* 23(3):513–52.
14. Read, S. 1995. *Thinking About Logic.* Oxford: Oxford University Press.
15. Restall, G. 2000. *An Introduction to Substructural Logics.* London: Routledge.
16. Restall, G. 2005. "Multiple Conclusions." In *Logic, Methodology and Philosophy of Sciences*, edited by P. Hajek, L. Valdes-Villanueva, and D. Westerståhl, 189–205. London: King's College Publications.
17. Rosset, J. V. 2006. *Some Logical Arguments Against Strict Finitism, Communication at the (Anti-)Realisms: Logic & Metaphysics Conference*. France: Nancy.
18. Salerno, J. 2000. "Revising the Logic of Logical Revision." *Philosophical Studies* 99:211–27.
19. Tennant, N. 1997. *The Taming of the True.* Oxford: Clarendon Press.
20. Wright, C. 1993. *Realism, Meaning and Truth* (2nd Edition). Oxford: Blackwell.

Chapter 4
Moore's Paradox as an Argument Against Anti-realism

Jon Cogburn

Phil.: How say you, Hylas, can you see a thing
which is at the same time unseen?
Hyl.: No, that were a contradiction.
Phil.: Is it not as great a contradiction to talk
of conceiving a thing which is unconceived?
Hyl.: It is.
Phil.: The tree or house therefore which you think of,
is conceived by you.
Hyl.: How should it be otherwise?
Phil.: Without question, that which is conceived is in the mind.
Phil.: How then came you to say, you conceived a house or
tree existing independent and out of all minds whatsoever?
George Berkeley—"Three Dialogues Between Hylas and
Philonous" [1].

And the criterion of a conceptual scheme different from our
own now becomes: largely true but untranslatable.
The question whether this is a useful criterion
is just the question how well we understand the notion of truth,
as applied to language, independent of the notion of
translation.
The answer is, I think, that we do not understand it
independently at all
Donald Davidson—"On the Very Idea of a Conceptual
Scheme" [6]

4.1 Introduction

At least since the time of Socrates, philosophy's ur-problem has been the proper account of the relation between truth and human epistemic capacities. For Socrates' heirs (henceforth "realists"), truth is primary in the sense that correct belief is to be

J. Cogburn (✉)
Louisiana State University Baton Rouge, Baton Rouge, LA, USA
e-mail: joncogburn@yahoo.com

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_4,

explained in terms of an impersonally conceived truth. For heirs of Socrates' sophist opponents (henceforth "anti-realists"), this is reversed. Truth itself is a kind of correct belief, which of course must be characterized non-circularly, without reference to truth.

If first-order logic seems like the natural, undisputed core of logic, this is because it is prima facie invariant between the realist and anti-realist orders of explanation. While establishing this point is old-going for most of us, it is still worth being explicit about.

On the dominant realist conception of consequence, a set of sentences Γ entails a sentence α if, and only if, it is not logically possible for all of the sentences in Γ to be true while α is false. Here consequence is made sense of via a (for the realist, impersonal) conception of truth. On the anti-realist conception, Γ entails α if, and only if, there is a proof of α from Γ. Since proofs are human constructions, consequence is made sense of in terms of human epistemic capacities.

Of course, these two conceptions are only as realist and anti-realist as their central concepts. It is always possible for the anti-realist to agree with the realist's conception of logical consequence while giving an anti-realist account of the central notions of logical possibility and truth.[1] Likewise the realist can argue that the anti-realist conception is right so far as it goes, but hold that the notion of "logical proof" can only be made sense of via the notion of an impersonally conceived truth.

When we think of things this way we see that cursory examination of the rhetoric of logic illustrates a strong realist bias. A proof system is standardly called *unsound* if it allows proofs not justified by the realist conception of consequence. A proof system is called *incomplete* if it does not permit construction of proofs for every case of realist consequence.

For example, consider the claim that second order logic is incomplete. By this we mean that any attempted second order consequence relation that is axiomatizable (in the sense that the set of proofs is checkable by a finitely statable procedure) will fail to provide means to prove cases of realistically acceptable consequence. This raises two interesting questions:

1. how did this terminology arise, and
2. are there any good reasons for it?

Why isn't what we universally call incompleteness of anti-realist consequence rather called the unsoundness of the realist consequence?

It is clear that proof theory is the poor relation here, either succeeding or failing to measure up to the semantic notion of consequence. Is this rational, or mere realist prejudice?

[1] See [14] for the most sophisticated, and I think promising, approach to this strategy.

4.2 Moorean Validity and Proof Theoretic Semantics

There are many ways to assess our problem. For example, Dummett struggles with what an anti-realist should say about Gödel's incompleteness results. Since saying the relevant Gödel sentences are false is not an option, Dummett [9] puts forward an open-textured (non-axiomatizable) notion of proof by which the Gödel sentence is provable.[2] Neil Tennant [15] has articulated non-question-begging proof theoretic criteria for logicality that prevent rules such as Prior's "Tonk" [12], which has the introduction rule of disjunction and the elimination rule for conjunction.[3] One might say that Tonk is bad because it allows derivation of falsehoods from truth, since it allows one to infer anything from anything. P entails $(P\ tonk\ Q)$, and $(P\ tonk\ Q)$ entails Q. As Tennant realizes in motivating his own proof-theoretic notion of logicality, the anti-realist cannot appeal to this fact in prohibiting Tonk, because to do so would be to appeal to the realist's notion of logicality.

Interestingly, Moore's Paradox[4] raises many of the same issues as Tonk. A sentence exhibits Moorean paradoxicality if it is true yet such that any sincere utterance of that claim is false. So consider any proposition P of which I have not thought. Even if it is true that P and it is true that I do not believe that P, I cannot sincerely utter "P and I do not believe that P". But if "P and I do not believe that P" is contradictory, it would seem that I believe all true sentences.

If there are contextually determined contradictory statements in Moore's sense there are certainly also contextually determined tautological statements, such "It is not the case that, (P and I do not believe that P)". But then we have the recipe for what might be called Moorean inferences. Consider the following.

P; therefore, P is believed.
P; therefore a being capable of asserting P exists.
P is provable; therefore P is proven.
P is provable; therefore P is known.

In each case, showing the antecedent to be fulfilled requires showing the consequent to be the case. If I show of an individual P that it is provable, I have proven P.

Of course we all know that there are provable propositions that have not yet been proven. We also know that it is logically possible for a proposition to be true without anyone, at least in practice, being capable of asserting it. Common sense says that Moorean validity is not real validity, and here common sense seems to be strongly on the side of the realist! It seems that the reason we know Moorean validities to be in fact invalid is because we know that it is possible for the premises to be true and

---

[2] The anti-realist needs to say something like this for second-order logic as well, which I think is probably the beginning of an explanation for our use of the soundness and completeness terminology, which apply to axiomatizable logics, albeit this issue needs to be investigated in light of the proposals in [14].

[3] For an argument to the conclusion that the quantified version of Tonk rather surprisingly undermines Christopher Peacocke's anti-Dummettian views, see [4].

[4] Moore is reputed to have raised this in a lecture.

the conclusion false, even though we can always transform a proof of the premises of a Moorean valid inference into a proof of the conclusion.

On reflection, the anti-realist should begin to feel a sinking feeling in her heart. Consider the following demonstration.

1. Moorean validities are not valid.
2. The realist explanation of why Moorean validities are invalid is better than an anti-realist explanation.
3. Therefore the realist order of explanation is correct.

Since premise 2. is the most contentious one, assessing this argument requires presenting the anti-realist semi-formal semantics of logical operators.

Following Dummett, contemporary anti-realists utilize Heyting, or proof theoretic, semantics to recursively correlate verification conditions with sentences of formal languages. For sentences bound with the logical operators of first order logic the correlations can be given in this manner.

Definition 1 (Heyting Semantics Definition of Truth[5]) Φ is true if and only if there exists a proof k of Φ, where one inductively defines what it is for k to be a proof of Φ as follows:

1. If Φ is atomic, $\Phi = (\Phi_1\alpha_1, \ldots, \alpha_n)$ where Φ_1 is an n-ary predicate, and each α_i is either an individual variable or constant, then k proves Φ if, and only if, k yields a general method that determines for which n-tuples of objects, $(< \beta_1, \ldots, \beta_n >)$, $(\Phi_1\alpha_1, \ldots, \alpha_n)$ holds.
2. If Φ is a conjunction, $\Phi = (\Phi_1 \& \Phi_2)$, then k proves Φ precisely when k yields a general method that enables us to find a proof k_1 of Φ_1 and k_2 of Φ_2.
3. If Φ is an implication, $\Phi = (\Phi_1 \rightarrow \Phi_2)$, then k proves Φ precisely when k yields a general method that, from every construction l_1 proving Φ_1, enables us to find a construction l_2 proving Φ_2.
4. If Φ is a negation, $\Phi = (\neg\Phi_1)$, then k proves Φ precisely when k proves $(\Phi_1 \rightarrow \perp)$, where the constant $\perp$ has no proof verifying it.
5. If Φ is a disjunction, $\Phi = (\Phi_1 \vee \Phi_2)$, then k proves Φ precisely when one can extract from k information about which of the terms Φ_i of the disjunction is true and a construction k_i proving that term Φ_i.
6. If Φ is an existential, $\Phi = (\exists x \Phi_1(x))$, then k proves Φ_1 precisely when k enables us to determine for which object a, $\Phi_1(a)$ holds.[6]
7. If Φ is a universal, $\Phi = (\forall x \Phi_1(x))$, then k proves Φ_1 precisely when k yields a general method that, for every object a, enables us to find a proof k_a of the proposition $\Phi_1(a)$.

Then, consequence is understood in the same terms as the conditional; Γ entails α if, and only if, every proof of Γ can be transformed into a proof of α. Importantly, these

[5] These clauses have been taken, with minor modifications, from [7].

[6] Utilizing these clauses, of course, requires standard use of alphabetic variants.

are semi-formal clauses. Anti-realists have developed a variety of formal proof-theoretic approaches to do justice to them.[7]

For our purposes, it is important to see how the clause for conditional clearly seems to license Moorean inferences as being valid. For example, any proof that P is provable can obviously be transformed into a proof that P is proven. The proof that P is provable will be a proof of P. Since Moorean inferences simply are such inferences that context of utterance solicits the transformation, the anti-realist is in real trouble.

4.3 On the Inadvisability of Biting the Bullett

An extreme anti-realist[8] could attempt to bite the bullet here, rejecting premise 1. of the previous argument. Consider the claim that the Moorean inference concerning provability and knowledge is false. That is:

$$\exists P((\vdash P) \wedge \neg KP)$$

Given the definition of the existential by Heyting Semantics, this seems like an absurd thing to say. It would say that we possess a procedure v for finding a particular P, such that v proves P and proves that P is not known. This seems obviously self-contradictory. So the hope is that, given the meaning of the anti-realist's words, there is no problem with affirming the Moorean validities as valid.

Unfortunately, this biting of the bullet does not work. If the existential binds propositions of a decidable theory, anti-realistically acceptable proofs exist for the above sentence. Take a complete theory such as Presburger's Arithmetic. Since it is provably (by intuitionistic means) complete we have the first premise of the following proof.

4.3.1 Antirealists Should Reject Unrestricted Moorean Validity

Consider a sentence of Presburger's Arithmetic not yet proven, such that we don't know yet if it is true or false. Call this sentence Q. Then by the following intuitionistically valid inference we obtain the unwanted conclusion.

---

[7] See [7, 14].

[8] In the body of this text I stick to the orthodox view (called by Tennant [15] *passive* manifestationism) that Dummettian anti-realists only charge understanders with being able to recognize canonical verifications of propositions when presented with them. For *this* position it is easy to establish the inadvisability of biting the bullet. However, in [8] the author argues for a more radical anti-realism where the understander must actually be able to *discover* the verifier (this, called by Tennant *active* manifestationism). Such a view is much more congenial to the unrestricted validity of Moorean inferences. Dubucs argues that his position requires further revision than mere intuitionism. If biting the bullet is more plausible for the feasibilist anti-realist, then it might be a test case for feasibilist revision that it prohibit the validity of the arguments in this section.

(1)	$\forall P(\vert\!-_{\mathrm{PA}}P \vee \vert\!-_{\mathrm{PA}}\neg P)$	
(2)	$\neg \mathrm{K}Q \wedge \neg \mathrm{K}\neg Q$	
(3)	$(\vert\!-_{\mathrm{PA}}Q \vee \vert\!-_{\mathrm{PA}}\neg Q)$	1 $\forall$ elimination
(4)	$\vert\underline{(\vert\!-_{PA}Q)}$	assumption for $\vee$ elimination
(5)	$\vert\neg \mathrm{K}Q$	2 $\wedge$ elimination
(6)	$\vert\vert\!-_{\mathrm{PA}}Q \wedge \neg \mathrm{K}Q$	4, 5 $\wedge$ introduction
(7)	$\vert\exists P(\vert\!-_{\mathrm{PA}}P \wedge \neg \mathrm{K}P)$	6 $\exists$ introduction
(8)	$\vert\underline{(\vert\!-_{\mathrm{PA}}\neg Q)}$	assumption for $\vee$ elimination
(9)	$\vert\neg \mathrm{K}\neg Q$	2 $\wedge$ elimination
(10)	$\vert(\vert\!-_{\mathrm{PA}}\neg Q \wedge \neg \mathrm{K}\neg Q)$	8, 9 $\wedge$ introduction
(11)	$\vert\exists P(\vert\!-_{\mathrm{PA}}P \wedge \neg \mathrm{K}P)$	10 $\exists$ introduction
(12)	$\exists P(\vert\!-_{\mathrm{PA}}P \wedge \neg \mathrm{K}P)$	3, $4-7$, $8-10$ $\vee$ elimination[9]

Therefore, the anti-realist must admit that there is something wrong with Moorean validity, even though Moorean validity seems to be valid by the anti-realist's own logical semantics! This is deeply paradoxical.

The anti-realist cannot forgo premise 1. of the previous section's master argument. The above argument shows that, for the intuitionist anti-realist at least, Moorean validities are not unrestrictedly valid. Then, since the realist can explain why these are not valid, and the anti-realist cannot, anti-realism is wrong.

4.4 A New Restriction Strategy

In the remainder of this paper, I will begin to sketch a solution for the anti-realist, as well as how this solution lends itself to a critique of the work of a surprisingly wide range of philosophers. In light of the above discussion, the main tasks for the anti-realist are:

1. to give a non-question-begging proof-theoretic restriction that prohibits arguments to absurdity involving Moorean inferences, and
2. to show that this restriction is reasonable in light of Heyting Semantics.

---

[9] At first blush, this argument does not run afoul of the revisionary strictures of Dubucs' feasibilism. Nothing is assumed for vacuous discharge, and nothing assumed for further discharge is used more than once inside the scope of that thing's subproof. However, a radical feasibilist might refuse to assert that a sentence is provable or disprovable just because that sentence is in a decidable theory. For it is such an assertion that forces an intuitionist to countenance proofs such that it is not feasible that any human could construct them.

For the first step, consider the following argument to the conclusion that $\exists P(\vdash_{PA} P \wedge \neg KP)$ is inconsistent with the Moorean inference from the provability of P to the claim that P is known.

4.4.1 Proof That i's Conclusion Is Inconsistent with Unrestricted Moorean Validity

1	$\exists P(\vdash_{PA} P \wedge \neg KP)$	proved in previous proof
2	$\quad (\vdash_{PA} P \wedge \neg KP)$	assumption for $\exists$ elimination
3	$\quad (\vdash_{PA} P)$	2, $\wedge$ elimination
4	$\quad \neg KP$	2, $\wedge$ elimination
5	$\quad KP$	3 Moorean validity
6	$\quad \bot$	4, 5, $\bot$ introduction
7	$\bot$	1, 2 − 6, $\exists$ elimination

Since every step in the previous proof and in this one (with the exception of step 5) is intuitionistically valid, the reasonable solution is to think the anti-realist must reject Moorean validity. But notice something interesting about this proof. The Moorean inference in question ($\vdash P$; therefore KP) is applied inside the scope of something assumed for further discharge. If the anti-realist can motivate the restriction of Moorean inferences as not applying in the scope of things assumed hypothetically, then she can argue that $\exists P(\vdash_{PA} P \wedge \neg KP)$ is consistent with (the now restricted) Moorean validity.

That this is a reasonable restriction for the classicist as well is shown by the following proof.

4.4.2 The Classicist Also Needs the Proposed Restriction

1	$\quad\quad (\vdash P)$	assumption for $\rightarrow$ introduction
2	$\quad\quad KP$	by Moorean inference
3	$\quad (\vdash P) \rightarrow KP$	1 − 2, $\rightarrow$ introduction
4	$\quad \forall P((\vdash P) \rightarrow KP)$	1 − 3, $\forall$ introduction

Allowing Moorean inferences inside the scope of things hypothetically assumed allows one to deduce things that are, for the realist, transparently absurd.

Now here is the main issue. The realist obviously has the semantic wherewithal to explain why $\forall P((\vdash P) \rightarrow KP)$ is an absurd claim, and why we shouldn't allow Moorean inferences to be applied to claims assumed hypothetically. Does the anti-realist? The intuitionistically acceptable proof above of $\exists P(\vdash_{PA} P \wedge \neg KP)$ does entail that $\forall P((\vdash P) \rightarrow KP)$ is false. Moreover, the proof involved propositions in a provably complete theory, propositions which we do in some sense possess the means of proving, but which (not having proven them yet) we don't currently know. Such theories are exactly where an intuitionist would expect $\forall P((\vdash P) \rightarrow KP)$ to fail. The realist, on the other hand, might think that in some sense a proof of P is in Plato's heaven, without ever being accessible to beings like us.

So we do have genuine proof theoretic means to show why Moorean validities are not always valid. Using them inside the scope of assumptions hypothetically assumed for further discharge leads to things that are absurd by the anti-realist's own lights. But since this condition is a proof-theoretic one, the anti-realist can use it to restrict the application of Moorean validities.

That being said, at this point one should still worry how the restriction accords with Heyting Semantics understandings of the operators, as well as whether the restriction is ad hoc. Nonetheless, at this point we do seem to have a promising start. The anti-realist notes that proofs to contradictions are generated only when Moorean validities are allowed inside of subproofs that rest on premises hypothetically entertained. This proof-theoretic analysis of the problem admits a proof-theoretic solution. Banning such application is an entirely proof-theoretic way to cope with the problem. Moreover, since the restriction applies to the meaning of deictic words such as "believes that" and "knows that" and "is proven", the solution keeps in place the Heyting Semantics account of the standard logical operators.

The charge of ad hocness is of course much more subtle.[10] In the remainder of this paper, I will begin to address this issue by showing how this restriction ramifies out in extremely non-trivial ways.

4.5 Is Antirealism a Moorean Validity? Reflections on Fitch's Proof and Dummett's Program

Contemporary anti-realism of the Dummettian sort rests upon verificationism, in our notation represented by the following inference

$$V.: P \vdash \Diamond KP$$

This says that P implies that it is possible to know that P.

Unfortunately, Frederic Fitch [11] showed that $V.$ entails the manifestly absurd claim that all truths are known. In light of our discussion of Moorean validity, it is worth re-examining Fitch's reasoning. His proof presupposes the validity of one

[10] See [3] for a sustained defense of Tennant's [15] solution to Fitch's paradox as not being ad hoc. Analogous arguments can be made in the case of the present solution.

transformation and four inference rules concerning knowledge and necessity. The transformation rule can be presented in this manner.

$$\Box - intro \ : \vdash P, \text{ therefore } \vdash \Box P$$

This rule says that if P is provable from no assumptions, then the necessity of P is also provable from no assumptions. Three of the other inference rules are:

$$K - elim.: \ KP \vdash P$$
$$K \wedge -dist.: \ K(P \wedge Q)) \vdash KP \wedge KQ$$
$$\Box\neg - dist.: \ \Box\neg P \vdash \neg\Diamond P$$

The first of these says that if there exists someone who knows a claim, then that claim is true; the second that if a conjunction is known then each of its conjuncts is also known; and the third that anything that is necessarily not the case is also impossible. With these rules, we can present Fitch's proof in the following manner.

4.5.1 Fitch Style Proof of Fitch's Paradox

1	$K(P \wedge \neg KP)$	assumption for $\neg$ introduction
2	$KP \wedge K(\neg KP)$	1, $K\wedge$ distribution
3	KP	2, $\wedge$ elimination
4	$K(\neg KP)$	2, $\wedge$ elimination
5	$\neg KP$	4, K elimination
6	$\bot$	3, 5, $\neg$ elimination
7	$\neg K(P \wedge \neg KP)$	1, 6, $\neg$ introduction
8	$\Box\neg K(P \wedge \neg KP)$	7, $\Box$ introduction
9	$\neg\Diamond K(P \wedge \neg KP)$	9, $\Box$, $\neg$ distribution
10	$P \wedge \neg KP$	assumption for $\neg$ introduction
11	$\Diamond K(P \wedge \neg KP)$	10, , V.
12	$\bot$	9, 11, $\neg$ elimination
13	$\neg(P \wedge \neg KP)$	10, 12, $\neg$ introduction
14	$P \rightarrow KP$	13 (strictly classical)[11] equivalence

[11] Intuitionists might reject the transition from line 13. to 14., as $P \rightarrow KP$ only really follows from $\neg(P \wedge \neg KP)$ with the help of a classical negation rule such as the law of excluded middle or

An extensive literature has risen up around this proof,[12] with most articles canvassing possible ways an anti-realist can block it. Our reaction to Moorean validities suggests a new strategy for Fitch's proof. Perhaps the anti-realist's inference $V.$ is itself a Moorean validity, and as such subject to the restriction that it not apply within the scope of things hypothetically assumed. But then Fitch's proof is invalid. In the above formulation, line $11.$ would not follow from line 10.

Given the problems with Fitch's proof, if this restriction is reasonable, it would be a tremendous boon to the anti-realist. Unfortunately, reflection on Dummett's program shows this to be disastrous. The contemporary semantic anti-realist program has a negative aspect and a positive aspect. Negatively, verificationism is used to criticize classical semantics and metaphysical positions that can be shown to in some way presuppose such semantics. Positively, classical truth conditions are to be replaced with something more constructively acceptable such as proof conditions. The important questions for any "restriction" strategy to Fitch's proof are the following.

1. Is the strategy well-motivated or ad hoc?
2. Is the resulting restricted $V.$ principle sufficient for Dummett's negative program? Can it work as a premise in plausible criticisms of classical semantics?
3. Is the resulting restricted $V.$ principle sufficient for Dummett's positive program?[13]

Does the kind of verificationism mentioned yield verification conditions that will suitably replace classical truth conditions? While it is not possible to answer all of these in this paper, I think that it is clear that the restricted version $V.$ is insufficiently weak to motivate Dummett's negative program.

The most plausible arguments for intuitionistic revision[14] paradigmatically consider a hypothetical sentence that by the classicist's account of the meaning of the logical connectives, ends up being (at least possibly) true yet unknowable. Then $V.$ is invoked to show that the sentence in question would, since it is true, be knowable. Then the contradiction is used to reject the classicist construal of the meaning of the logical vocabulary. Clearly, if $V.$ could not be applied inside the scope of something assumed for further discharge, such arguments would be invalid.

This being said, it should be noted that the classicist should feel free to accept $V.$ restricted as above. Asserting P sincerely commits one to being able to assert that it is possible to know that P. Dummett's argument for revision shows that the classicist had better not think that a hypothetical assertion of P allows one to conclude that it is hypothetically possible to know that P. So while we haven't solved Fitch's paradox, the line between the realist and anti-realist can be marked in an interesting way.

double negation elimination. In [16], the author suggests that this might be thought of as providing evidence for intuitionism, albeit not very much. The denial that any claim can be both true and unknown, as stated schematically in line (13), is problematic enough.

12 See the discussion in [3].

13 Tennant [15] conclusively shows that Dummett's own restriction strategy does not yield a verificationism strong enough to motivate the Dummettian program.

14 See [5] for my regimentation of Dummett's argument.

4.6 Further Reflections on Fitch's Proof

If our restriction hasn't deflected Fitch's proof, it does save Tennant's deflection of Fitch's proof. Tennant [15] argues that $V.$ should be restricted to Cartesian principles, that is, principles which are not logically inconsistent to assume known. Thus, Tennant's restricted principle is:

$$V.' : \neg(KP \vdash \bot),\ \text{therefore}\ P \vdash \Diamond KP$$

To see how this blocks derivation of Fitch's paradox, note that in the above derivation, line 10. is arrived at by applying $V.$ to $P \wedge \neg KP$. But lines 1. through 6. show $K(P \wedge \neg KP)$ to be logically inconsistent. So by Tennant's revised verificationist principle, line 10. in Fitch's proof does not follow.[15]

The most recent paper in the literature on Tennant's revision is Berit Brogaard and Joe Salerno's "Clues to the Paradox of Knowability" [2]. While Brogaard and Salerno's masterful exploration of modal notions in light of Fitch's proof sheds very bright light on a number of issues,[16] their critique of Tennant is as wrong as wrong can be. Brogaard and Salerno do not present their proof in a natural deduction system. Rather, it is a numbered sequence of propositions, each motivated in natural language. Given this, scope issues are never made explicit. Once one fully regiments their argument, it is clear that they use a Moorean inference in an unacceptable way.[17] In addition to Tennant's restricted verification principle, the only novel inferences they use are:

$$(*)\ \text{If}\ K\neg A,\ \text{then}\ \neg\Diamond A,$$
$$P \vdash KP.$$

The first says that if A is known to be false then A is not epistemically possible. The second is the Moorean inference in question. Here is the fully regimented proof.

---

15 Again, see [3] for an argument that Tennant's solution is not ad hoc.

16 Arguably, the greatest contemporary task facing semantic anti-realists is discerning a plausible anti-realistically acceptable theory of modality. Timothy Williamson's [17] paper remains canonical. For further philosophical discussions, see especially [2, 13], and Salerno's dissertation. My own view, still being worked out in light of Dummettian dialectical pressures and reasonable success conditions on any theory of modality, is a development of the Humean idea that impossibility is projected on the world via our own experience of not being able to fulfill all of our desires, and that "objective" possibility and necessity are parasitic on this projection. Unlike Hume, but like Schopenhauer, I take it that phenomenology shows us that the experience of frustrated desire is objective. I also do not think that this renders modal claims about the universe, math, and logic to be devoid of truth values, albeit bivalence does fail.

17 In what follows, the following correspondences of lines with numbered propositions in Brogaard and Salerno's proof hold: lines 1–5 with propositions 1–5; lines 6–9 with propositions 5.1–5.4; lines 10–21 with propositions 6–6.92; lines 22–23 with propositions 7–8; lines 24–26 with propositions 9–9.2; lines 27–28 with propositions 10–11. In footnotes I give their justification for each numbered proposition.

4.6.1 A Regimentation of Brogaard and Salerno's Argument Against Tennant

1	$\exists A(\neg KA \wedge \neg K\neg A)$	assumption for $\neg$ introduction[18]
2	$\neg KA \wedge \neg K\neg A$	assumption for $\exists$ elimination[19]
3	$\diamond K(\neg KA \wedge \neg K\neg A)$	2, $V.'$[20]
4	$K(\neg KA \wedge \neg K\neg A)$	assumption for $\neg$ introduction[21]
5	$K\neg KA \wedge K\neg K\neg A$	4, $K\wedge$ distribution[22]
6	$K\neg KA$	5, $\wedge$ elimination
7	$\neg\Diamond KA$	6, $*$
8	$K\neg K\neg A$	5, $\wedge$ elimination
9	$\neg\Diamond K\neg A$	8, $*$
10	$\neg\Diamond KA \wedge \neg\Diamond K\neg A$	7, 9, $\wedge$ introduction[23]
11	A	assumption for $\neg$ introduction
12	$\Diamond KA$	11 by $V.'$
13	$\neg\Diamond KA$	10, $\wedge$ elimination
14	$\perp$	12, 13, $\perp$ introduction
15	$\neg A$	11 – 14, $\neg$ introduction
16	$\neg A$	assumption for $\neg$ introduction
17	$\Diamond K\neg A$	16 by $V.'$
18	$\neg\Diamond K\neg A$	10, $\wedge$ elimination
19	$\perp$	17, 18, $\perp$ introduction

[18] "*Let us suppose (for our primary reductio) that there is an undecided statement:*" [2, p. 147].

[19] "*If line 1 is true, then some instance of it is true:*" [2, p. 147].

[20] "*Since line* 2 *does not violate Tennant's restriction (that is,* $K(\neg KA \wedge \neg K\neg A)$ *is not self-contradictory), we may apply anti-realism to it. It follows from anti-realism that it is possible to know* 2*:*" [2, p. 147].

[21] "*Now let the anti-realist suppose for reductio that it is known that A is undecided.*" [2, p. 147].

[22] "*Knowing a conjunction entails knowing each of the conjuncts. Therefore,*" [2, p. 147].

[23] "*Applying principle* $(*)$ *to each of the conjuncts gives us*" [2, p. 147].

20	$\neg\neg A$	$16-19$, $\neg$ introduction
21	$\perp$	$15-20$, $\perp$ introduction
22	$\neg K(\neg KA \wedge \neg K\neg A)$	$4--21$, $\neg$ introduction[24]
23	$K\neg K(\neg KA \wedge \neg K\neg A)$	22 Moorean validity[25]
24	$\neg\Diamond K(\neg KA \wedge \neg K\neg A)$	23, $*$[26]
25	$\perp$	$3-24$, $\perp$ introduction
26	$\perp$	$2-25$, $\exists$ elimination
27	$\neg\exists A(\neg KA \wedge \neg K\neg A)$	$1-26$, $\neg$ introduction[27]
28	$\Box\neg\exists A(\neg KA \wedge \neg K\neg A)$	27, $\Box$ introduction[28]

Note that, once we fully regiment the proof, it is clear that the application of the Moorean Principle occurs inside the scope of two things assumed for further discharge, lines 1 and 2 of the proof, negation and existential respectively. If the above proof were valid, Tennant would be in serious trouble, as his restricted form of verificationism would entail that it is necessarily the case that there does not exist a sentence that we don't know to be true or false. However, the proof is not valid. Consider the following proof to the same conclusion.

[24] "*Given the assumption of anti-realism, we derive a contradiction* $(\neg A \wedge \neg\neg A)$. *So the anti-realist must reject our assumption at line* 4." [2, p. 148]. Note that one can craft a shorter proof if one just deletes line 16, allows lines 17 to follow from 15 by $V.'$, removes the discharge bars from 17 and 18 deletes 19 and 20, and has absurdity in 21 follow from 17 and 18. That is, starting another subproof on line 16 is superfluous, as we already have $\neg A$ on line 15. I've presented it with the superfluous subproof only to make clear that I really am regimenting the proof Brogaard and Salerno had in mind.

[25] "*Resting only on the assumption of anti-realism, which the anti-realist takes to be known, line* 7 *[22] is now known*" [2, p. 148].

[26] "*But then, by* $(*)$, *it is epistemically impossible to know that A is undecided:*" [2, p. 148].

[27] "*But this contradicts line* 3, *which rests merely upon anti-realism and line* 2. *Line* 2 *is the instance of the undecidedness claim at line* 1. *A contradiction then rests on anti-realism conjoined with undecidedness. The anti-realist must reject the claim of undecidedness:*" [2, p. 148].

[28] "*Since anti-realism is taken to be a necessary thesis, it must be admitted by the anti-realist that, necessarily, there are no undecided statements:*" [2, p. 148].

4.6.2 The Same Argument Without Tennant's Principle

1	$\exists A(\neg KA \wedge \neg K \neg A)$	assumption for $\neg$ introduction
2	$\neg KA \wedge \neg K \neg A$	assumption for $\exists$ elimination
3	A	assumption for $\neg$ introduction
4	KA	3 Moorean validity
5	$\neg KA$	2, $\wedge$ elimination
6	$\perp$	4, 5, $\perp$ introduction
7	$\neg A$	3, 6, $\neg$ introduction
8	$K \neg A$	7 Moorean validity
9	$\neg K \neg A$	2, $\wedge$ elimination
10	$\perp$	8, 9$\perp$ introduction
11	$\perp$	$2-10$, $\exists$ elimination
12	$\neg \exists A(\neg KA \wedge \neg K \neg A)$	$1-11$, $\neg$ introduction
13	$\Box \neg \exists A(\neg KA \wedge \neg K \neg A)$	10, $\Box$ introduction

This shows that Brogaard and Salerno's use of the Moorean validity inside the scope of things assumed for further discharge allows one to prove the undesirable consequence *without the use of anti-realism*. Verificationism is superfluous in their proof! So they have not produced a counterexample to Tennant's theory, but have rather just provided roundabout evidence for the correctness of my proposed restriction on the application of Moorean validities.

If the restriction cannot be used to disarm Fitch's proof, at least it disarms an influential criticism of Tennant's anti-realist solution to Fitch's proof.

4.7 Berkeley and Davidson's Use of Moorean Validities

In closing, I'd like to suggest that other philosophers use Moorean inferences in the proscribed way. In this light, it is interesting to note that Dummett [10] seems to suggest that his form of anti-realism is the post-linguistic-turn descendent of Berkeleyan idealism.[29] For Dummett, the philosophy of language gets at some of the literal content of the metaphysical picture presented by Berkeley. Interestingly, one can argue that Berkeley himself uses Moorean reasoning in his defense of idealism. In the quote that opened this paper, Philonous asks Hylas to try to conceive of

[29] One could argue, on historical and philosophical grounds, that Dummettian anti-realism is really the most plausible current form of neo-Kantianism.

something existing yet unconceived. Since one can't do this, and since conception is a kind of perception, it follows that to exist is to be perceived. Unfortunately for the idealist, our restricted version of Moorean validities renders this argument invalid. Hylas is concluding about a hypothetically presumed object that it is conceived. This clearly runs afoul of our proscription.

Perhaps more contentiously, one can argue that Donald Davidson utilizes a Moorean inference very similar to Berkeley's in the above quote. In *On the Very Idea of a Conceptual Scheme*, Davidson asks us to imagine a language not translatable into ours. But since to imagine a language would be to imagine it as being translatable into ours, we can't do this. Therefore every language is translatable into ours, and thus (along with other Davidsonian premises) there are no radically different conceptual schemes. Like Berkeley, he is using a Moorean inference inside the scope of something assumed hypothetically. We are to conclude about a hypothetically presumed language, that since we have a conception of it, it must be translatable into our language. While this may true of any actual language we can so categorize (for the very Davidsonian reason that we would not be in a position to consider it a language unless we could translate it into ours) to conclude this about a language hypothetically assumed to exist requires improper use of Moorean reasoning.

These are contentious claims, but at the very worst, defenders of Berkeley and Davidson, as well as critics of Tennant, must now attempt to explicate the relevant arguments in a way that does not run afoul of our restriction. If not agreement, at least further understanding will be achieved.

Acknowledgements I would like to thank Berit Brogaard, Emily Beck Cogburn, Jacques Dubucs, Jeff Roland, Joe Salerno, Mark Silcox, and Neil Tennant for conversation and inspiration. This paper arose out of sustained reflection on [2, 8]. I would have liked to further discuss the parts of Brogaard and Salerno's piece that strike me as both non-trivial and true (i.e. the important restrictions on various options for an anti-realist account of modality), as well as to have explored a Dubucsian reaction to Moorean validities in more detail. However, space-time constraints, combined with the internal logic of the paper, prohibited this.

References

1. Berkeley, G. 2005. *Three Dialogues Between Hylas and Philonous*. London: Longman.
2. Brogaard, B., and J. Salerno. 2002. "Clues to the Paradoxes of Knowability." *Analysis* 62: 143–50.
3. Cogburn, J. 2004. "Paradox Lost." *Canadian Journal of Philosophy* 34(2):195–216.
4. Cogburn, J. 2005a. "Tonking a Theory of Content: An Inferentialist Rejoinder." *Logic and Logical Philosophy* 13:31–36.
5. Cogburn, J. 2005b. "The Logic of Logical Revision: Formalizing Dummett's Argument." *The Australasian Journal of Philosophy* 83(1):15–32.
6. Davidson, D. 1982. "On the Very Idea of a Conceptual Scheme." In *Inquiries into Truth and Interpretation*, 183–99. Oxford: Oxford University Press.
7. Dragalin, A. 1980. *Mathematical Intuitionism, Introduction to Proof Theory*. Providence, RI: American Mathematical Society.
8. Dubucs, J. 2002. "Feasibility in Logic." *Synthese* 132:213–37.

9. Dummett, M. 1978. "On the Philosophical Significance of Gödel's Theorem." In *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press.
10. Dummett, M. 1991. *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.
11. Fitch, F. 1963. "A Logical Analysis of Some Value Concepts." *The Journal of Symbolic Logic* 28:135–42.
12. Prior, A. 1960. "The Runabout Inference-Ticket." *Analysis* 21:38–39.
13. Salerno, J. 2000. "Revising the Logic of Logical Revision." *Philosophical Studies* 99:211–27.
14. Tennant, N. 1986. "The Withering Away of Formal Semantics?" *Mind and Language* 1: 302–18.
15. Tennant, N. 2002. *The Taming of the True*. Oxford: Oxford University Press.
16. Williamson, T. 1987. "On the Paradox of Knowability." *Mind* 96:256–61.
17. Williamson, T. 1992. "On Intuitionistic Modal Epistemic Logic." *Journal of Philosophical Logic* 21:63–89.

Chapter 5
The Neutrality of Truth in the Debate Realism vs. Anti-realism

María J. Frápolli

5.1 Introduction

There is an essential aspect of Ramsey's account of truth that has been systematically neglected: his use of the term 'prosentence' to explain how truth ascriptions work (vid. [12]). An exception has been Engel and Dokic's book [11]. Ramsey's awareness of the fact that it is easy to understand what truth is, the real difficulty being to *say* what is is surprising. His explanation of the fact, that natural languages do not have enough expressions able to play the role that is played in artificial languages by propositional variables is even more surprising. This is an essential role, by the way, one that cannot be dispensed with. My aim here is not historical, though. The Ramseyian insight has been developed independently of Ramsey's works by some philosophers before and after him and credit to them will be paid below in the appropriate places. My concern here is systematic, and it also has a practical derivation. The systematic part is to offer a sketch of an enriched prosentential account of truth. It is a sketch because a completely thorough presentation would require too much material for a paper, although this sketchy presentation will, I hope, convey enough information so as to tempt the reader to move towards the theory. It is enriched because it pays attention to syntactical aspects, semantic contributions, and pragmatic roles. In the end, the enriched view will have the virtue of placing together several ideas that proceed from different approaches to truth, and show how they can co-exist in a consistent and powerful proposal.

The practical derivation is related to the place of truth in the debate between realism and antirealism. I will say it directly: *none*. The truth predicate plays a variety of different tasks in natural languages, all of them essential to their expressive power, but both our comprehension of truth and the use we make of the truth predicate are strictly independent of our *theories* about the relation between mind and world.

M.J. Frápolli (✉)
Departamento de Filosofía I, Universidad de Granada, Granada, Spain
e-mail: frapolli@ugr.es

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_5,

5.2 Truth

The truth predicate works as a builder of prosentences. Prosentences are the natural language equivalent of propositional variables in artificial languages. An exhaustive account of the meaning of truth in natural languages can be offered by way of explaining the syntactic, semantic and pragmatic roles performed by the truth predicate, following the threefold traditional distinction, due probably to Peirce and recovered by Morris. Let's state the theory broadly:

A. The syntactic job of the truth predicate is restoring sentencehood.
B. A sentence that has truth as its main predicate is a *truth ascription* and truth ascriptions are proforms of the propositional kind, i.e. *pro-sentences*. The semantic role of prosentences, as that of the rest of proforms, is threefold: they work

   i as vehicles of direct propositional reference,
   ii as vehicles of anaphoric reference, and
   iii as instruments for propositional generalization.

C. Finally, the pragmatic role of truth ascriptions is the endorsement of propositional contents, i.e. the explicit acceptance of propositional contents as ready to be used in inferential exchanges.

Prior's [20] and Horwich's [18] characterization of the truth operator as a denominalizer and also Quine's disquotationalism focus upon the syntactic role of the truth predicate as a mechanism of restoring an expression's syntactical category of SENTENCE.

What is currently known as 'the prosentential view' stresses the semantic purpose of truth ascriptions. Truth ascriptions are prosentences and prosentences are a special kind of proform. Proforms, as natural language variables, are dummy expressions that reproduce the role of any instance of the logical category they belong to. Pronouns are the best known among proforms, but they are not the only ones. Proadjectives, proadverbs, and prosentences are also proforms, and natural languages host many expressions that work as these not-so-well-known auxiliary expressions. When linguists qualify an expression as a 'pro-noun' they classify it in the category of singular terms. Indeed, a pronoun is a term that can be substituted by any singular term *salva gramatica*. Nonetheless, the perspective taken here is different, since we are classifying expressions according to their logico-semantic behaviour rather than according to their syntactic status. Some expressions that function as pro-nouns from a syntactic point of view turn out to be pro-adverbs, pro-adjectives or even pro-sentences whenever they are considered from a logical point of view. Words like 'it' and 'that' can inherit any content whatsoever, and are thus all-purpose (or *transcendental*) proforms. This will become clear in what follows.

The credit of the term 'prosentential theory' has to be given to several people that originally employed it without having any knowledge of its use by others. Bolzano was the first philosopher to use the expression '*Fürsatz*'[1] with the meaning that we give to the term 'prosentence' here and, as Ramsey [21] did some years later, he attributed the status of prosentences to the grammatical adverbs 'yes' and 'no'. Seventy years after Bolzano's use and almost 50 years after Ramsey's, Grover, Camp and Belnap [16], on the one hand, and Williams [26], on the other, developed the prosentential account independently.

The pragmatic ingredient of the enriched account presented here is not new either. Pragmatically oriented philosophers of language have recognized the pragmatic role of truth ascriptions in the act of endorsing a content. Strawson [25] offered a pragmatic view on truth in which the truth predicate works as a marker of illocutionary force. Nevertheless, Strawson's view cannot be reduced to this claim. Besides stressing its role as a force marker, Strawson recognizes other roles of the semantic notion *par excelence*. In his paper 'Truth', Strawson says:

> In many of the cases in which we are doing something besides merely stating that X is Y, we are available, for use in suitable contexts, certain abbreviatory devices which enable us to state that X is Y [...] *without* using the sentence-pattern 'X is Y'. Thus, if someone asks us 'Is X Y?', we may state (in the way of denial) that X is not Y, by saying 'It is not' or by saying 'That's not true'; [...]. It seems to me plain that in these cases 'true' and 'not true' (we rarely use 'false') are functioning as abbreviatory statement-devices of the same general kind as the other quoted.[2]

The British philosopher takes the truth operator to be a way of codifying ranges of statements and, in his view, it is neither exclusively a force marker nor a redundant expression. A few lines below the text quoted above, Strawson says:

> It will be clear that, in common with Mr. Austin, I reject the thesis that the phrase 'is true' is logically superfluous, together with the thesis that to say that a proposition is true is *just* to assert it and to say that it is false is *just* to assert its contradictory. 'True' and 'not true' have jobs on their own to do, *some*, but by no means all, of which I have characterized above.[3]

This is a crucial remark, for to say that an expression has a particular pragmatic significance doesn't preclude its eventual semantic meaning and its syntactic function.

Recently, Robert Brandom [3] has insisted upon the pragmatic role of truth ascriptions. Truth, Brandom maintains, helps to make the commitments and entitlements of our claims explicit. A truth ascription displays the speaker's endorsement of a propositional content. By qualifying a propositional content as true, the speaker commits herself to that content as something for which she is ready to give reasons, if required. By accepting that content as true, one is giving permission to use it as a premise in further inferential acts.

I endorse the semantic core of the prosentential theory of truth and propose completing it with the syntactic insights given by Prior, Quine, and Horwich, on the one hand, and with the pragmatic picture developed by Strawson and Brandom, on the other. Taking all this information into account, a comprehensive theory can be

---

[1] See [2]. I owe this information and the reference in Bolzano to Göran Sundholm to whom I am deeply grateful.

[2] [25, pp. 174–75].

[3] [25, p. ivi].

concocted of how the truth operator works, i.e. a theory that explains its inferential behaviour, that answers the essential philosophical questions traditionally related to truth, and that serves as the point of departure of the declaration of independence of truth from metaphysical and epistemic disputes which is one of the main aims of this paper.

5.3 Realism and Antirealism

The realism/antirealism debate comes in (at least) two flavours: metaphysical and epistemic. The semantic formulation of the debate due to Dummett, who defines realism as related to classes of statements rather than to classes of entities, is reducible to one of the two.[4] The debate is patent in the philosophical disputes between the different proposals about the notion of truth. There are theories of truth that explain truth as a metaphysical notion (correspondence to facts), and some others that explain it in epistemic terms (the coherence of one's belief system, assertibility, etc.), and it is not uncommon that the realism/antirealism debate turns into the correspondence/coherence debate or into the truth vs. assertibility debate.

Metaphysical realism states the independence of reality from our thought and will. A realist statement about a particular domain (metaphysics, ethics, aesthetics, semantics, logic) is the acknowledgement of the existence of facts of the appropriate kind, i.e. it is the acknowledgement of the existence of metaphysical facts, moral facts, aesthetical facts, semantic facts, or logical facts. Once the existence of the appropriate kind of fact is assumed, truth is standardly defined as correspondence with facts of the kind in question. Truth is ascribed to a proposition if there is a fact that makes the sentence true. This fact is sometimes known as the sentence's truth-maker.

Epistemic realism, in turn, states the objectivity of knowledge. Since knowledge is traditionally understood as justified true belief, the notions of truth, knowledge and objectivity allegedly lie on the realist's side. Antirealism is then left with the task of defining diluted substitutes for these central concepts because, the classical story goes, there is no room in an antirealist context for robust notions of truth, objectivity, or knowledge. This is the standard view, and the view that I will challenge.

Truth is neither a metaphysical nor an epistemic notion, as Tarski has already claimed, and a complete account of truth able to explain the meaning and use of a truth operator is compatible with any particular position in metaphysics and epistemology. The debate between realists and antirealists doubtless raises profound philosophical questions, but none of the parties are justified in claiming exclusive rights on truth, knowledge and objectivity. Truth is generally involved in metaphysical and epistemic debates partially at least because the truth operator is

[4] See Dummett, [8, p. 56] and [9, p. 564]. Semantic realism is not an independent brand. It relies either on metaphysical realism or on epistemological realism, depending on the way in which one assumes that meaning and content are reached at.

an indispensable instrument of propositional generalization, and metaphysical and epistemic discourse are classical contexts in which we deal with general thoughts.

Truth ascriptions play their role once some propositional contents have been accepted. The home of the realism/antirealism debate is the justificatory level, i.e. how and why we assume that some contents are claimable or, to put it another way, the dispute between realists and antirealists emerges in relation to the question of how to accept the truth-maker itself, i.e. the content of the truth ascription. Only afterwards the truth predicate appears in the picture. This point is particularly relevant for the realism/antirealism debate, for it shows that there can be a neutral definition of truth that both parties, realists and antirealists, are allowed to use. Besides, removing the question of truth from the metaphysical and epistemic discussion allows us to sort out some the specific difficulties related to the definition of truth in natural languages and some others concerning the structure of reality and our access to it.

5.4 The Prosentential View

An account of truth is called 'prosentential' if it interprets the truth operator as a means of forming natural language pro-sentences. A pro-sentence is a pro-form of the sentential kind, i.e. a sort of propositional variable. A welcome consequence of prosententialism is that it considers the truth predicate as a member of a general kind, the kind of proform builders. It shows that the notion of truth is not resistant to analysis, that a definition of it can be offered for natural languages, and that it is possible to explain the role it performs while avoiding the two extreme views of considering it either primitive, and hence indefinable, or else trivial, and therefore also indefinable.

5.4.1 The Semantic Functions of the Truth Predicate

Let's begin with semantics since the semantic analysis of truth constructions has been the trademark of prosententialism. Typically, pro-forms perform three semantic tasks: they are vehicles of (α) direct reference, (β) anaphoric reference, and (γ) generalization. Since most of our everyday universal quantifiers are binary operators, i.e. operators that need two concepts to construe a complete proposition, nearly all cases of (γ) are also cases of (β). Let us consider some examples.

A. Pronouns

a.1 *This* is my car
a.2 I heard about this car and I bought *it*
a.3 If I own a car, I take care of *it*
[$\forall x(x$ is a car & I own $x \rightarrow$ I take care of $x)$]

These three are examples of pronouns working as cases of (α) – (a.1), (β) – (a.2) and (γ) – (a. 3). In (a.3), the pronoun 'it', and the last variable 'x' in its logical form, are bound variables that permit generalization, and at the same time they are anaphorically linked to their heads, 'a car' in the natural language example, and the value of the first variable 'x' in the antecedent of the conditional, in the semi-formalized case.

Natural languages also contain pro-adverbs, pro-adjectives and pro-sentences. Most natural language expressions performing pro-adverbial, pro-adjectival and pro-sentential functions are not included into the grammatical category of adverbs, adjectives and sentences respectively. A difficulty that the prosentential view has to face is that natural languages paradigmatically use pro-nouns, i.e. expressions with the syntactic category of singular terms, to perform the logical roles of the rest of pro-forms.

B. Proadverbs

The following examples contain pro-adverbs:

b.1 I love being *here*
b.2 I will go to Miami and will be *there* till Christmas
b.3 Every*where* I go, I meet nice people *there*
[$\forall l$(I go to $l \rightarrow$ I meet nice people in l)]

Again, (b.1) is a case of pro-adverb in a direct referential use, (b.2) is a case of pro-adverb in an anaphoric referential use, whose head is 'Miami', and (b.3) is a case of pro-adverb performing a generalization function (and anaphoric reference).

C. Proadjectives

The following are examples of pro-adjectives:

c.1 What colour will you paint the house? I would like my house to be *this* colour [pointing at a sample]
c.2 Granada used to be parochial, but now it is not *so*.
c.3 Victoria is some*thing* that Joan is not (*so*)
[$\exists v$(Victoria is v& Joan is not v)]

In (c.1), 'this' functions as a pro-adjective replacing a colour word. In (c.2) 'so' works as a variable that anaphorically refers to the adjective 'parochial', and in (c.3) 'something' is a quantifier that ranges over qualities, so that the instances of (c.3) have to include adjectives in the argument place.

That there are pro-forms other than pronouns in natural languages is something that has been widely recognized. A mere glimpse of Ramsey, Prior, Grover, and Williams will be enough. If we are convinced that the class of pro-forms is wider than the class of pro-nouns, then the acknowledgement of pro-sentences should be almost routine.

D. Prosentences

Pro-sentences are typical pro-forms, and as such they perform the same three tasks performed by the rest of pro-forms. Let us see some examples:

d.1 What did she say? She said *this* [pointing to a sentence in a newspaper]
d.2 Zapatero said that *peace was close* and Rajoy denied *it*
d.3 Every*thing* President Obama says is ratified by Hilary Clinton
[$\forall p$(President Obama says that $p \rightarrow$ Hilary Clinton says that p)]

In examples (d.1)–(d.3), 'this', and 'it' have the syntactic category of pro-nouns, although the logical category of pro-sentences, and '-thing' in the quantifier also binds pro-sentences. A slight paraphrase of (d.3) will clarify this:

d.3* When President Obama says something, Hilary Clinton ratifies *it*.

There are some ready-made objections launched time and again against the analysis of pro-forms that we have put forward. The most 'obvious' is that this analysis requires higher-order quantification and that this obliges us to embrace an untenable ontology. First of all, proponents of the prosentential view are aware of this alleged obstacle, they just consider this objection untenable. There is no reason to maintain, *pace* Quine and his followers, that quantification exhibits our ontological commitments. In natural languages we use quantifiers related to all kind of expressions. We say that some skylines are more impressive than some others, that there are many ways of cooking rice, or that some of our most secret desires are hard to explain, without feeling that our ontology is overcrowded with skylines, ways of cooking rice, and secret desires together with our familiar medium size objects. And we are right. Ontology is signalled by referential expressions, and quantifiers and the variables bound by them are not of this kind.[5]

Using what has been said so far as theoretical background, let us now turn to the explanation of truth. Languages need pro-forms because they are the only means of anaphoric reference and generalization. Direct reference and the direct expression of a content can be achieved by proper names, in the case of reference to objects, and by genuine adjectives, adverbs or sentences, in the case of the non-mediated expression of a semantic value. But without proforms, i.e. without mechanisms for anaphora and generalization, the expressive power of languages would be considerably shortened. Some uses of pro-forms are acknowledgedly uses of laziness, but the vast majority of them are not; in cases of anaphoric reference and of genuine generalization[6] pro-forms cannot be dispensed with. Examples of pro-sentences used out of laziness are responsible for the widespread, false idea that the truth operator is

[5] To a highly convincing and deeply informed defence of non-nominal quantification see [20, 27].

[6] By a genuine generalization I understand one that is not equivalent to a finite conjunction.

redundant.[7] Cases of anaphoric reference and genuine generalization show why it is not. In general, the truth operator is as redundant as any other kind of pro-form, and we have independent theories that explain that pronouns and demonstratives are essential to the expression of some kinds of first-person thoughts,[8] cross references, and general contents.

E. Complex Prosentences

In a formal language such as that of propositional calculus we have single propositional variables, the sentential letters. In other formal languages, in the first order predicate calculus for instance, we can interpret formulae as complex propositional variables of a certain kind. Different formulae correspond to natural language sentences with different structures. Natural languages[9] possess the same variety of expressions. They have single propositional variables, although unfortunately, there are only two of them, 'yes' and 'no'. Unlike 'it', 'this', 'what' and others that can act as proforms of different categories, 'yes' and 'no' are the only natural language proforms that are essentially prosentences. Grammar characterizes 'yes' and 'no' as adverbs, but from a logical point of view the type of pro-form a particular token belongs to does not depend on its syntactic category but rather on the kind of item from which it inherits its content. In this case, 'yes' and 'no' inherit complete propositional contents. These two unique single propositional variables are patently not enough to do all the work that pro-sentences have to do. Nevertheless, natural languages have other resources. In particular, they have means of building up a wide diversity of complex propositional variables. Some of these means are the formal predicates 'is true', 'is a fact' and others. In the following examples, the definite description 'What he said is true' works as a complex prosentence that inherits the content of the previous sentence that acts as its anaphoric head:

e.1 He said that Americans are proud of their country. What he said is true
e.2 "Victoria never lies", said John. What he said is true

The content of the truth ascription in (e.1) is that Americans are proud of their country; the content of the truth ascription in (e.2) is that Victoria never lies. In both cases, the prosentence does not have a content in itself, but serves as a vehicle of any propositional content that is contextually salient. This is compatible with the fact that the prosentence doesn't change its meaning from an occasion of use to another one. The truth ascription is not ambiguous; its linguistic meaning, i.e. its character, remains constant. The fact that a truth ascription can change its content

[7] All proforms, prosentences included, have uses of laziness. The truth predicate has this use in all versions of the Tarskian T-sentences. This is the grain of truth behind the redundancy theory of truth.

[8] See for instance the explanation about quasi-indicators due to H–N. Castañeda [5, p. 74].

[9] We are referring to Indo-European languages, although it is not too risky to suppose that the use of variables of different categories is a semantic universal.

from context to context without changing the meaning of the truth predicate has motivated the spurious debate about whether there are different notions of truth, i.e. the monism vs. pluralism debate on truth. The notion of truth is univocal from the point of view of the linguistic meaning, although a truth ascription can acquire different contents depending on the item from which it inherits its content. The situation here is hardly more puzzling than the fact that that the pronoun 'he' can be used to refer to my son, to my father and to the King of Spain.

In examples (e.1) and (e.2) the prosentence is performing anaphoric references. In (e.3) and (e.4) they act as mechanism for propositional generalization:

e.3 Everything that follows from a true theory is true
e.4 Everything the Pope says is true

That the truth operator is not redundant in natural languages obviously follows from the fact that general propositions cannot be expressed without proforms, prosentences in this case, since proforms are the expressions that accompany quantifiers.

5.5 The Syntactic Function of the Truth Predicate

The truth predicate also performs an indispensable syntactic function. In the previous examples with the exception of those in the first group (a.1)–(a.3), the syntactic category of the pro-form does not coincide with its logical status. In (d.3)*, for instance, the expression that is a pro-sentence from a logical point of view has the status of a pro-noun. Nevertheless, there are situations that require pro-sentences to possess the syntactic status of sentences. That is, there are situations in which a pro-sentential use of, say, 'it' needs to be supplemented to become an expression with the syntactic status of a well-formed sentence to preserve the rules of grammar.

Imagine that Victoria utters 'I do not like Mondays' to express the proposition that she does not like Mondays. We can refer to her claim by different means. We can say that she really believed what she said, and here 'what she said' is the prosentence. When we refer to a proposition, we use an expression appropriate for referring, i.e. a singular term, and in these cases what is logically a pro-sentence is syntactically either a pro-noun or a definite description. A useful way of referring to propositional contents in the written language is using inverted commas.[10]

In the same way in which natural languages have mechanisms to squeeze complete propositions into singular terms (the use of syntactic pro-nouns as prosentences), they also have mechanisms to execute the opposite movement, i.e. to

[10] Inverted commas have many other uses, not only this one, and when they are the mechanism of reference they do not always refer to a content. They can refer to the sentence itself, either type or token, or to some aspects of it. See, for instance, [4, 7, 17, 22, 24] for different accounts of the way in which inverted commas function.

unleash a prosentence codified in a pronoun into a complete sentence. If we call the former mechanism 'nominalizer', we can also call the latter mechanisms 'de-nominalizer'. Recall that this is the function that Horwich [18] concedes to the truth predicate, and it is a generalization of the famous Quinean disquotationalism. The two functions of obtaining singular terms out of propositions, on the one hand, and propositions out of singular terms, on the other, end in what the Kneales [19] have dubbed as 'designations of propositions' and 'expressions' of them, respectively. Let us consider an example

Proposition (expressed by Victoria's utterance 'I do not like Mondays'): Victoria does not like Mondays.

Designation of the proposition (*exhibitive*): 'Victoria does not like Mondays'.

Designation of the proposition (*blind*): What Victoria said.

Expression of the proposition (*exhibitive*): 'Victoria does not like Mondays' is a true sentence.

Expression of the proposition (*blind*): What Victoria said is true.

The terms 'exhibitive' and 'blind' are intended here to stress that in some truth ascriptions the anaphoric head from which it is possible to recover the content of the prosentence is exhibited in the very ascription, whereas there are cases (the *blind* ones) in which this does not occur. There are other denominalizers in natural languages. '... is a fact' is a well-known one, a false friend that has nurtured the correspondence theories of truth. 'What Victoria said is true' is a prosentence (or a prosentence and the dummy truth predicate, it depends on the authors[11]) constructed out of a blind designation of a proposition and a denominalizer. Its content is dependent on the content of its anaphoric antecedent, i.e. the proposition to which it is anaphorically linked. In the previous example its content is that Victoria does not like Mondays, but in different situations it can inherit any propositional content whatsoever. 'What Victoria said is a fact' has exactly the same structure and function, and thus connecting the two expressions (or their contents) by an equivalence sign results in a true claim, 'What Victoria said is true *iff* it is a fact', but that does not take us closer to the understanding of any of the predicables involved.

Thus, the syntactic function of the truth predicate is converting designations of propositions into expressions of them, restoring the status of sentencehood to singular terms that already have propositions as their contents. As a historical curiosity, Frege assigned in his *Begriffsschrift*[12] the same syntactic function to the formal predicate 'is a fact'. And his intuitions were correct: 'is true' and 'is a fact' are exactly the same type of operator, with the same range of syntactic and pragmatic functions.[13]

---

[11] Ramsey, Strawson, Horwich and Brandom offer a separate treatment of the truth predicate, while Grover, Camp and Belnap deal with complex pro-sentences like 'what he said is true' as a block.

[12] [14, p. 3].

[13] The semantic function of prosentences was completely alien to Frege's views.

5.6 The Pragmatic Function of the Truth Predicate

We aim at truth when we produce assertions, and both notions, *truth* and *assertion*, belong to the same family of notions, they need each other. They are interdefinable, although their interdefinibility simply means that we are characterizing a particular linguistic game of which they both are constitutive. The pragmatic task of truth is making some of our inferential commitments explicit. But what kind of commitment does a truth ascription make explicit? It makes explicit that we are engaged in a speech act with the force of a claim, although this is not its only task. Austin was accused by Strawson[14] of reducing the meaning of truth to this expressive role. Since it brings into the open the force of a claim as a claim, the truth predicate makes explicit the appropriateness of using its inherited content as something for which reasons can be given and demanded. In ascribing truth to a proposition we are disclosing our doxastic commitments to it.[15] A truth ascription explicitly identifies a content as something to be counted among the available information, ready to be used in our inferential games. This can be done either by welcoming a proposition into one's beliefs system for the first time or else by transferring contents from some circumstances, in which they have been accepted as claimable, to some other circumstances (considered sufficiently relevant as to permit a safe transfer).

Truth ascriptions by which we directly refer to a salient proposition, i.e. ascriptions of the 'it's true' type, are cases in which we allow the referred proposition to enter the system of accepted information. The status of accepted information is highly context-dependent, and a proposition can be so characterized for some purposes, and thus welcomed as true, while in some other circumstances, or for different purposes it can be rejected, and its entrance to the system vetoed. Once propositional contents have entered into the system of accepted knowledge, it is possible, using the truth operator, to generalize about them. But recall that the truth ascription does not produce nor cause the epistemic status of 'accepted knowledge'. It merely sanctions it, makes it explicit and, by means of the rest of logical notions, the truth operator permits to handle propositional contents and possibly reorganize and project the information as in the case of generalizations.

5.7 Epistemology and Metaphysics

Depending on the particular theory of justification one favours, the reasons for the acceptance of some content vary. One can accept a proposition because, say, one considers that it has been reached in the aftermath of a reliable process, or because it coheres with the rest of our beliefs, or because the scientific community acknowledges that it has passed the standard procedures of justification in the corresponding

[14] See [25, p. 182].

[15] Nowadays, Brandom [3] has put this notion of claim as something for which the speaker is responsible into the fore. The same insight is found in Frege [15, p. 281], where he contrasts assertion with what an actor does on stage.

discipline, or because the linguistic community at issue democratically accords its acceptability, and so on. This is the first step, the step that is subjected to epistemic discussion. The truth operator operates at a second stage, and it lies outside the epistemic discussion, i.e. it operates on the outputs of the justification processes. These processes can be positioned on any zone of the justificatory spectrum, they can be scientific procedures or assumptions of common sense, and they can be empirical or a priori, formal or informal. All this belongs to epistemology and pragmatics. And it is only subsequently that the result obtained by the epistemological processes will eventually be inherited by an explicit ascription of truth.

How linguistic or mental entities acquire content is another disputed subject, to which different theories offer different answers. The two wide paradigms that practically exhaust the spectrum are, at present, truth-conditional semantics, and its contextualist version, on the one hand, and inferential semantics, on the other.[16] At face value, truth-conditional semantics appears closer to metaphysical realism, whereas inferential semantics shows relevant points of contact with antirealism. Nevertheless, this impression is inaccurate. The core of a truth-conditional treatment of content is that the content of an utterance is its truth conditions. But this claim only means that the content of an utterance are the conditions under which it is true. What are the conditions under which Victoria's utterance of the sentence 'I don't like Mondays' is true? Obviously, that Victoria doesn't like Mondays. And what are the truth-conditions of the claim that through a point external to a straight line only passes one parallel? Well, that through a point external to a straight line only passes one parallel. What about the claim that water is H_2O? It will be true if, and only if, water is H_2O, and so on. But again, one can affirm that Victoria doesn't like Mondays, that for a point external to a straight line only passes one parallel, that water is H_2O, and so on both from a realist view about how the world is constituted and also from an antirealist perspective. The discussion depends on how we reach a position in which we are allowed to make these affirmations and on our general understanding about the relation between humans and their surroundings. Similarly, the theoretical core of inferential semantics amounts to saying that the content of a linguistic or mental act with the force of a claim are the contents from which it follows and the contents that follow from it, i.e. the application conditions, entitlements, and their consequences, their commitments. Both realists and antirealists agree on the set of contents from which it follows and those that follow from it. Thus, strictly speaking, the four possible combinations—truth-conditional semanticist and realist, truth-conditional semanticist and anti-realist, inferential semanticist and realist, inferential semanticist and anti-realist—are all legitimate. Truth-ascriptions are means of endorsing contents, contents that are sometimes displayed in the very ascription, and sometimes are not; contents that are sometimes singular and sometimes general, but the meaning of the truth predicate is independent of these features, and it is involved neither on the debates about content, nor on the debates about realism and antirealism. Nevertheless, it is unquestionable that

[16] An example of truth-conditional pragmatics is [23]; an example of inferential semantics is [3].

the notion of truth appears profusely in epistemic and metaphysical discussions, and justifiably so. Nevertheless, the justification is not that truth is either an epistemic or a metaphysical notion. It is not. The notion of truth is not conceptually involved in these debates but it is, so to say, put to the test. Let me briefly explain this last claim.

Although truth is not an epistemic notion, the truth predicate is omnipresent in epistemological discourse; and not even the most basic theses in epistemology can be stated without essentially using the truth predicate. Besides, the endorsement role that the truth predicate performs in natural languages is applied in many cases to the items coming out of the justificatory filters sanctioned by epistemology. The prosentential account explains thus the insight that traces a connection between truth and justification. Besides, since the truth operator is a means of forming prosentences, i.e. propositional variables, it (or any equivalent operator) has to be around whenever propositional generalizations are needed. The truth operator, according to this use of building up general sentences, is the natural language counterpart of propositional quantifiers and proposicional variables in artificial languages. Epistemology and the philosophy of science are paradigmatic contexts in which we deal with packs of propositions, and natural languages can only deal with general contents by means of propositional proforms, i.e. prosentences.

Truth is not a metaphysical notion either, although metaphysics is another context in which the use of prosentences is essential. The predicates 'is true' and 'is a fact' are both prosentence builders, and sentences like 'this is true' and 'this is a fact' are both propositional proforms. Being true, like being a fact, are natural language operators that convert singular terms, whose content is a complete proposition, into sentences; they also serve to construe both singular and general prosentences. It cannot be denied that something is true if, and only if, it is a fact. It cannot be denied because it is an instance of the principle of identity. As an instance of the principle of identity, it has no informational content, but the correspondentist slogan that truth is correspondence with facts is empty in a further sense; the two sentential arguments that accompany the equivalence operator, i.e. 'something is true' and 'it is a fact', are actually pro-sentences; they are not sentences that can be used in isolation to express a content, for they are proforms that need an antecedent, or a referent. In this sense, there is no contradiction in embracing an antirealist perspective in epistemology and metaphysics while accepting at the same time the T-schemes of the Tarskian theory of truth, or the Aristotelian dictum that to say of what it is that it is not is the false, and of what is that it is is true, or any standard formulation of the Correspondence theory. There is nothing wrong in saying that truth is correspondence with facts, and that something is true *iff* it is a fact. There is nothing wrong, although the correspondentist claim is neither an explanation nor a definition, it is merely a periphrasis. This situation explains why most people agree on the correspondentist slogan, and at the same time disagree on the details of a theory of correspondence. The slogan is tautological, but its implementations are not.

Mixing up the realism/anti-realism debate with the definition of truth is the effect of a poor understanding of the way in which the truth operator works. The realism/anti-realism debate unquestionably touches upon fundamental philosophical questions, but none that have any effect for a theory of truth.

My conclusion is that the realist has no exclusive rights on the notion of truth,[17] and that the antirealist concedes too much to his opponent by renouncing his own rights on this essential notion. The notion of truth can be completely defined in a self-contained theory as the prosentential view. The prosentential view explains how the truth operator works and why it is indispensable in contexts in which we focus on general claims. It also accounts for the cogent insights behind the correspondence theory of truth and the theory of truth as redundancy. Furthermore, it shows that truth is neutral between realism and antirealism. So far the realists practice of reclaiming truth for their cause has been extremely successful, but it is as unjustified as the antirealist renouncement of their proud use of it.

References

1. Blackburn, S., and K. Simmons, eds. 1999. *Truth*. Oxford: Oxford Readings in Philosophy, Oxford University Press.
2. Bolzano, B. 1904/1930. *Wahrheit und Evidenz* (2nd Edition, 1974). Felix Meiner Verlag.
3. Brandom, R. 1994. *Making It Explicit. Reasoning, Representing, and Discursive Commitment.* Cambridge, MA: Harvard University Press.
4. Cappelen, H., and E. Lepore. 1997. “Varieties of Quotation.” *Mind* 106:429–50.
5. Castañeda, H. N. 1967. “Indicators and Quasi-Indicators.” In *The Phenomeno–Logic of the I. Essays on Self-Consciousness*, edited by J. G. Hart and T. Kapitan, 61–88. Bloomington, IN: Indiana University Press.
6. Castañeda, H. N. 1999. *The Phenomeno–Logic of the I. Essays on Self-Consciousness*, edited by J. G. Hart and T. Kapitan. Bloomington, IN: Indiana University Press.
7. Davidson, D. 1984 [1979]. *Quotation. Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
8. Dummett, M. 1976. “What Is a Theory of Meaning? (II).” In *The Seas of Language,* edited by M. Dummett, 34–93. First published in *Truth and Meaning*, 1976, edited by Gareth Evans and John McDowell. Oxford: Oxford University Press.
9. Dummett, M. 1992. “Realism and Antirealism.” In *The Seas of Language*, edited by M. Dummett, 462–78. Oxford: Oxford University Press.
10. Dummett, M. 1993. *The Seas of Language*. Oxford: Oxford University Press.
11. Engel, P., and J. Dokic. 2003. *Frank Ramsey. Truth and Success*. London/New York, NY: Routledge.
12. Frápolli, M. J. 2005a. “Ramsey’s Theory of Truth and the Origins of the Prosentential Account.” In *F. P. Ramsey. Critical Reassessments*, edited by M. J. Frápolli, 113–38. London: Continuum.
13. Frápolli, M. J., ed. 2005b. *F. P. Ramsey. Critical Reassessments.* London: Continuum.
14. Frege, G. 1879/1952. “Begriffsschrift. A Formalized Language of Pure Thought Modeled Upon the Language of Arithmetic.” In *Translations from the Philosophical Writings of Gottlob Frege*, edited by P. Geach and M. Black, 1–20. Totowa, NJ: Barnes and Noble.
15. Frege, G. 1903. “Foundations of Geometry: First Series.” In *Collected Papers on Mathematics, Logic and Philosophy*, edited by Brian McGuinness, 273–84. Oxford: Basil Blackwell.

[17] With the notions of knowledge and objectivity the situation is similar, although an analysis of them lies outside the scope of this paper.

16. Grover, D., J. Camp, and N. Belnap. 1975. "A Prosentential Theory of Truth." *Philosophical Studies* 27:73–125.
17. Haack, S. 1974. "Mentioning Expressions." *Logique et Analyse* 17:67–8, 277–94.
18. Horwich, P. 1998. *Truth*. Oxford: Clarendon Press.
19. Kneale, W., and M. Kneale. 1962. *The Development of Logic*. Oxford: Clarendon Press.
20. Prior, A. 1971. *Objects of Thought*. Oxford: Clarendon Press.
21. Ramsey, F. 1929. "The Nature of Truth." In *On Truth*, edited by N. Rescher and U. Majer, 6–24. Original Manuscript Materials (1927–1929) from the Ramsey Collection at the University of Pittsburgh, Kluwer, 1991.
22. Recanati, F. 2000. *Oration Obliqua, Oration Recta. An Essay on Metarepresentation*. Cambridge, MA: The MIT Press.
23. Recanati, F. 2001. *Literal Meaning*. Cambridge, MA: Cambridge University Press.
24. Richard, M. 1986. "Quotation, Grammar, and Opacity." *Linguistic and Philosophy* 9: 383–403.
25. Strawson, P. F. 1950. "Truth." In *Truth*, edited by S. Blackburn and K. Simmons, 162–82. Oxford: Oxford Readings in Philosophy, Oxford University Press.
26. Williams, C. J. F. 1976. *What Is Truth?* Cambridge, MA: Cambridge University Press.
27. Williams, C. J. F. 1989. *Being, Identity, and Truth*. Oxford: Clarendon Press.

Chapter 6
Modalities Without Worlds

Reinhard Kahle

6.1 Modal Logic

As the standard *theory* for modalities, *modal logic* is nowadays well-established. It provides two new operators $\Box$ and $\Diamond$. With them 'ϕ is necessary' is represented by $\Box\phi$ and 'ϕ is possible' by $\Diamond\phi$. We assume that the reader is familiar with modal logic.[1] We just mention some of the basic properties which we discuss later on.

As one of the advantages of modal logic, often, the dualism of necessity and possibility is mentioned. One can introduce 'ϕ is possible' by 'not-ϕ is not necessary', i.e. $\Diamond\phi$ is an abbreviation of $\neg\Box\neg\phi$. Or, vice versa, $\Box\phi$ could be defined as $\neg\Diamond\neg\phi$. This is a very nice formal property. However, formal properties are not arguments for adequacy, and it would be a different task to argue for this duality on a philosophical level.

The core rule of modal logic is *necessitation*:

$$\frac{\vdash \phi}{\vdash \Box\phi}$$

Based on it, different modal logics are developed by adding particular axioms. For the popular system **S4** we have, for instance, the following ones:

K. $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$.
T. $\Box\phi \to \phi$.
4. $\Box\phi \to \Box\Box\phi$.

It is easy to see that these axioms serve only for an account of logical necessity (since there are no non-logical components involved). Here, we do not discuss the

[1] As standard reference we can refer to the introduction to modal logic by Bull and Segerberg in the *Handbook of Philosophical Logic, Volume II*, [3].

R. Kahle (✉)
CENTRIA and DM, FCT, Universidade Nova de Lisboa, P-2829-516 Caparica, Portugal
e-mail: kahle@mat.uc.pt

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_6, © Springer Science+Business Media B.V. 2012

adequacy of these axioms for logical necessity. We just mention that a good part of the philosophical discussion of modal logic is about the validity of certain principles. This is probably best exemplified in the controversy of the principle: 'What is possible, is necessarily possible.' It can be expressed in the following axiom which turns **S4** into **S5**:

5. $\Diamond\phi \to \Box\Diamond\phi$.

Observe that this is a principle involving nested modalities.

The formal systems of modal logic go already back to C. I. Lewis. Besides Łukasiewicz, it was in the first place Carnap who made modern modal logic popular, [4].[2] However, modal logic was considered incomplete as long as there was no *semantics* for it. This problem was resolved with the invention of *possible worlds semantics*.[3]

6.2 Possible Worlds Semantics

Possible worlds semantics was introduced independently by Kanger, Hintikka, and Kripke in the late 1950s.[4] As for modal logic we assume that the reader is familiar with the basic idea of possible worlds semantics.

It extends the well-known truth value semantics of Tarski for propositional or first-order logic by relativizing truth to *worlds*. In general, a proposition can be true in one world and false in another. Going back to Leibniz, the naive idea is that necessity is explained by 'truth in all possible worlds'. But, to deal with nested modalities, in addition, an *accessibility* relation between worlds is needed. With it, a proposition is necessary in a particular world, if it is true in all worlds accessible from there.

The possible worlds semantics exhibits an extraordinary technical elegance. For instance, the question about the adequateness of the axiom **5** turns into a question about algebraic properties of the accessibility relation (which, to verify **5**, would have to be *Euclidean*[5]).

Due to its virtues, possible worlds semantics is nowadays considered not only as the standard semantics for necessity, but it is also used as a fundamental tool to study other intensional phenomena like relevance or counterfactuals.

But possible worlds semantics comes for an enormous *philosophical* price only: The ontology explodes. Next to the actual world, one needs additional *possible*

---

[2] For a brief history of modal logic, see the first part of [3].

[3] In fact, Carnap's approach to modalities makes already extensive use of possible worlds; however, he did not provide a semantics for modal logic, since the accessibility relation was still missing.

[4] For a historical survey on possible worlds semantics, cf. [6]. An overview of the current philosophical discussion is given in [7].

[5] Therefore, this axiom is sometimes also designated by **E**.

worlds to interpret modalities. There are some philosophers, most prominently David Lewis, who see no problem in such an exploding ontology and advocate a modal realism which takes the different possible worlds as real.[6] We will not go into a discussion of this realism, but we like to raise just one question: Which meaning of *realism* is used here? Our problem is that the different possible worlds are per se disconnected: One cannot jump from one possible world to another. Thus, the objects in a possible world different from the one we live in are unreachable; in other words, their *realness* is—from our perspective—only hypothetical. But why these objects should deserve the attribute *real*? There is an obvious etymological relation between *real* and *realize*,[7] which is violated here. Albeit etymology does not provide directly philosophical arguments, we think that the realism of possible worlds semantics contorts the word *realism*. In fact, many philosophers[8] are more careful and use possible worlds only as a theoretical construction without conceding that they are *real*, cf. e.g., [19, p. 24]. But we will even go a step further and question the usefulness of possible worlds semantics at all. In fact, we would like to ask what possible worlds semantics actually provides *as a semantics*.

6.3 The Role of Semantics

Semantics should provide meaning. This can be achieved, for instance, for a term in a formal theory or an expression of natural language by an interpretation of this term or expression in another realm. Implicitly, it is assumed that this realm is already understood (or at least *better* understood as the formal language or the expression in the natural language).

From this point of view, possible worlds semantics offers probably a better understanding of the operators $\Box$ and $\Diamond$ than one can grasp directly from modal logic. But this leaves the question open whether modal logic, in fact, provides the proper formalizations of the notion of necessity and possibility. We doubt that it does.

Let us see how semantics usually works. When we consider standard first-order logic, the Tarskian semantics works a follows: For a first-order sentence, you have to provide the interpretation of the non-logical symbols, and the semantics returns you the truth value of the sentence. In analogy, possible worlds semantics works according to the principle: You have to give it the possible worlds (together with an accessibility relation), and it returns you the truth value of your modality statement.

But there is a substantial difference in the way the ‘additional’ ontology comes into play: for first-order logic we need some ontology to interpret the non-logical symbols. But this additional ontology is justified by the sheer use of these symbols; for modal logic the ontology of the possible worlds ‘appears from nowhere’. In

[6] Cf. [17]. For another discussion of the realism of possible worlds, see [5].

[7] This relation is not accidental; compare the German words *Wirklichkeit* and *bewirken*.

[8] They are dubbed *ersatzers* by David Lewis and his followers.

fact, we have absolutely no control over the ontology which slips into the variety of possible worlds in question.

It seems to us, that in a possible worlds semantics we have to ‘put in’ more than we ‘get out’: Neither necessity nor possibility statements involve any component addressing directly any ontological addition to it. But to obtain a truth value for such a statement we would be committed to such an addition. We would expect a semantics to be ontologically modest.

There is another point worth mentioning with respect to the *logic* we use: In first-order logic, we can easily restrict the semantics of the non-logic symbols by adding further (non-logical) axioms *involving these symbols*—this is somehow what first-order logic serves for. But how would look like the corresponding axioms restricting the variety of possible worlds? Since there are no symbols referring to *worlds* we can only speak about them implicitly; this makes the use of worlds *as semantics* doubtful.

6.4 Criticism of Modal Logic

In this section we ask in what way modal logic can be considered as a formalization of necessity and possibility.

We like to point out here a triviality which is usually not noted in an introduction to modal logic: If we do not consider nested operators, modal logic does not provide more than a box in front of the derivable formulas. Thus, the power of modal logic is located *only* in the nesting of operators.[9] We think, that statements with nested operators should not be the primary target of a theory for necessity and possibility; we would like to have first a complete understanding of the non-nested cases. But, to put just a box in front of formulas is a rather weak explanation of necessity.

In addition, the necessitation rule makes it impossible to speak within (ordinary) modal logic of necessarily true and contingently true sentences at the same time. With respect to this last point the situation is a little bit better in possible worlds semantics: Here, we can distinguish one world as the actual (or ‘real’) one.[10] And we should do so. But, on the modal logic side, this would require to add an extra *actuality operator*. Such extensions exist in the literature,[11] and although they are an

[9] Of course, syntactically one can speak about expressions like $\neg\Box\phi$. However, it seems to be the case that the standard systems of modal logic do not derive any interesting $\neg\Box\phi$ theorems. We heard that this criticism was repeatedly put forward by Quine, but we are not aware of any reference.

[10] And both, Leibniz and Carnap, make extensive use of the actual world in their approaches to possible worlds.

[11] See the list of references in footnote 9 of [27, p. 199]. In fact, Wehmeier proposes a modal logic in [26, 27] which is a special case in that it makes do without an actuality operator and instead employs a distinction between indicative and subjunctive predicate symbols.

important topic in the philosophical discussion they did not find their way into the standard presentations of modal logic. In any case, this addition does not overcome the main problem, that, if we restrict ourselves to the non-nested case, modal logic does not provide much more than a box in front of derivable formulas from a background theory.

Of course, one might consider it as the task of modal logic, to study such a background theory. We could, for instance, add certain axioms which represent metaphysical necessities, and so modal logic provides a framework to formalize metaphysical necessity in general (and if we do not add any non-logical axioms we just get a theory for logical necessity.) But, again, if we do not consider nested necessity statements, it is not clear what such a theory provides more than the plain derivability from these axioms.

And this is a criticism which carries over to the possible worlds semantics: The necessary truths are those statements which hold in all possible world. But if the variety of the possible worlds is fixed *first*, the semantics does not provide any new insights with respect to necessary statements. All what we get is a closure under some modal laws for nested necessity statements.[12]

Another problem of this account is logical omniscience. As modal logic is closed under logical equivalence, every possible world is closed under logical consequence.[13] This turns every logical truth into a necessary one. This is, of course, in accordance with the intuition for *logical necessity* (but not necessarily for other forms of necessity). But, we like to question whether the study of (this form of) logical necessity is of much interest; but even if, it is obviously not needed to use any modal argument to defend (or find) logical truths. So, what does modal logic serve for in this case?

In fact, also outside of logic, we are not aware of any practical examples where modal logic or possible worlds semantics helps us to determine a necessary truth which was not already (explicitly or implicitly) built in by certain axioms or constraints on the variety of worlds. To state it again, all we get out of these approaches is some kind of logical relations for nested modalities. And this is not surprising if one recalls the axioms of **S4**, for instance.

Note that we do not criticize the use of modal logic, and with it possible worlds as its (formal) semantics, *as formal systems*. But we question its usefulness in analyzing *necessity* and *possibility* in an informative way.

---

[12] This was observed, for instance, by Føllesdal in [8, p. 572f.]. Stalnaker [25, p. 333] is half on this way when saying: 'The possible worlds representation of content and modality should be regarded, not as a proposed solution to the metaphysical problem of the nature of modal truth, but as a framework for articulating and sharpening the problem.' We think that the contribution of possible worlds to articulate and sharpening the problem of modalities is rather limited.

[13] There are some attempts in the literature to 'define' worlds which are not closed under logical consequence, cf. e.g., [19]; however, we consider these approaches as rather premature. In particular, they are usually discussed only informally.

6.5 An Alternative Analysis of Modalities: Possibility

In the following we present an alternative analysis of modalities. In this analysis the modalities are considered as *metatheoretical* properties of sentences. 'Metatheoretical' means that we assume that a certain axiomatic framework is given, and the modalities are interpreted on the meta-level of this framework.[14]

6.5.1 Possibility as Independence

We start with *possibility* which is, in some sense, the easier case. Let us look to a generic example of the form:

(1) ϕ is possible.

Here, ϕ should be an arbitrary first-order sentence,[15] not involving further modalities. But ϕ should belong to a language for which we have an axiom system given, representing our state of discourse. This system plays an important role in the following and we call it $\mathfrak{A}$.[16]

Now, (1) can easily be explained as

ϕ is independent from $\mathfrak{A}$.

That means, neither ϕ nor its negation $\neg\phi$ is derivable in $\mathfrak{A}$.

This approach is by no means original; however, the reading of *possibility as independence* is here proposed as an alternative to the $\Diamond$ operator in modal logic, and we have to defend it in close comparison.

First let us remark that with our reading, possibility is of the same quality as provability. Therefore, we are on another logical *level* than in modal logic. This step to the meta-level, excludes, for the time being, the nesting of modalities. As remarked above, in modal logic non-nested necessity is essentially equivalent to derivability from the background theory. Therefore, a non-nested possibility statement $\Diamond\phi$, i.e., $\neg\Box\neg\phi$, could be expressed by the underivability of $\neg\phi$. Obviously, the independence of ϕ implies the underivability of $\neg\phi$. In this sense the notion of possibility proposed here is contained in the notion given in modal logic. However,

[14] We are not sure whether in this analysis necessity and possibility can still be classified as *modi* of sentences. If not, it would be odd to call them *modalities*. For this reason, it might be more correct to put quotes around *modalities* in the title: 'Modalities' without worlds.

[15] The restriction to a first-order language has the only purpose to fix a logical framework; the given analysis works for any other formal language of arbitrary complexity.

[16] However, the concrete formal calculus—a Hilbert-style calculus, a sequent calculus, natural deduction, or any other calculus—is not important.

we demand something more: the underivability of ϕ. This can expressed in the debatable property[17]:

The possibility of ϕ entails the possibility of $\neg\phi$ (and vice versa).

As consequence we get:

What is provable is not possible.

It might sound less odd, if we put it in the following form[18]:

What is provable is no longer possible.

This principle can be defended on *pragmatic* grounds: If we already know that something is the case, why should we classify it in the 'weaker' term of possibility. This pragmatic restriction of possibility can also be seen as a consequence of a Grice-like *maxim*, asking to avoid unnecessary weakenings of available information.[19] We are in favor of this principle, but confess that it is controversial. If you strongly reject it, you can adjust our analysis of (1) by reading it just ad '$\neg\phi$ is not derivable in $\mathfrak{A}$'.[20] In this way, it matches with the understanding of (non-nested) possibility in modal logic.

6.5.2 *Epistemic Possibility*

The principle has a side effect which helps with the problem of logical omniscience: Tautologies, which are derivable by definition, are automatically outside of the scope

---

[17] We have an ally of highest authority here: Aristotle derives this property from (one of) his definitions of possibility in the Prior Analytics, cf. [24, §5.6.3].

[18] We will come back the temporal aspect later, when we argue that possibility statements are normally statements about the future. A further specification of our principle could even be:

What is provable is no longer considered as possible.

The subjective aspect of 'considered' is in particular of relevance for the *epistemic modalities* discussed below.

[19] As far as we see, it does not follow from any of the original maxims of Grice [12]. In the form: 'Do not make your contribution less informative than necessary' it would be somehow a dual to the Maxim of Quantity: 'Do not make your contribution more informative than is required.'

[20] This matches, in an epistemic context, with the first reading of possibility by Frege in his *Begriffsschrift*: 'Wenn ein Satz als möglich hingestellt wird, so enthält sich der Sprechende entweder des Urtheils, indem er andeutet, dass ihm keine Gesetze bekannt seien, aus denen die Verneinung folgen würde; ...' [10, p. 5] ('Is a proposition is advanced as possible, either the speaker is suspending judgment by suggesting that he knows no laws from wich the negation of the proposition would follow ...', [11, p. 13]).

of possibility. We consider this as a virtue of our approach. However, it seems to be that we would accept statements like:

Riemann's Hypothesis is possible, as well as
The negation of Riemann's Hypothesis is possible.

But, we would put them in the form of:

It is possible that Riemann's Hypothesis is true, as well as
It is possible that the negation of Riemann's Hypothesis is true.

The part 'is true' can be taken—somehow contrary to its literal meaning—as an indication that these statements are not statements about Riemann's Hypothesis directly, but about *our knowledge* of it (or its proof/disproof). We would like to call this form of modality an *epistemic modality*. And, for epistemic modality, we can replace the *underiv**ability*** of ϕ and $\neg\phi$ (which corresponds to the *independence* of ϕ) by the fact the neither ϕ nor $\neg\phi$ is *actually deriv**ed*** (*so far*).[21] In this view, probably most of the possibility statements can be considered as epistemic ones.

6.5.3 The Future

It is essential for our account of possibility that the axiom system $\mathfrak{A}$ is not (syntactically) complete; otherwise, there would be no possible sentences left. For an axiom system describing our real world, there is a specific class of sentences predestined for possibility: Sentences about the future. In fact, sentences about the past are *not* candidates for possibility. Let us illustrate this by a modification of a classical example:

It is possible that there is a see battle yesterday.

This sentence has to be considered as odd, because of the tense of (the second) 'is'. We can, of course, express it in the past tense:

It is possible that there was a see battle yesterday.

But in this case, believing in the determinateness of the past, we can consider this only as a *epistemic possibility*; i.e., we do not have sufficient 'axioms' about yesterday to derive whether there was a see battle or not; the axiom system which

[21] A similar point was discussed in [13] in relation with contradictory *belief sets*; it might well be that we have contradictory beliefs, but as long as a contradiction was not *derived*, we can live happily with them.

describes the 'actual world' contains all this information, and there is no room any longer for possibility statements about the past.

Back to the future: It is not the case, that every sentence about the future is possible, as long as we allow axioms which determine the future.[22] Natural candidates for such axioms are *laws of nature*. As simple example we consider the 'law':

Every morning the sun rises.

For our analysis, it is not important whether this law is indeed true, but only that it makes part of the axiom system $\mathfrak{A}$ which *is taken* as a description of the real world. In this case a sentence like 'It is possible that the sun does not rise tomorrow morning.' turns out to be wrong.

We will not discuss the different possibilities (or impossibilities) of axioms speaking about the future in a description of the actual world.[23] It is just worth noting that the aspect of the future tense in possibility statements should be taken into account. If we turn once more to modal realism, this position would require to attribute *reality* to the future—in fact, to all possible futures. When we criticized that the variety of possible worlds is hard to grasp, this criticism becomes even more substantial if we think of possible worlds in the future.

6.5.4 Ontological Modesty

We like to emphasize that our analysis of possibility statements works *without* any reference to an additional ontology; it works on a purely syntactic level, the same level on which derivability is settled. Since a syntactically incomplete axiom system allows for different interpretations, we can consider every such interpretation as a *world* in the sense of possible worlds semantics.[24] And it should be straightforward how these worlds, containing different facts about the future, but respecting our axioms, look like. And we can even interpret our possibility statement now as true in one of these worlds. But, this is just *secondary*! We can do so, but do not have to do so. Possible worlds are possible but not necessary for the analysis. They are an epiphenomenon. In particular, the philosophical question of the realism towards possible worlds is secondary.

Here it is worth citing a remark of Saul Kripke from his famous lectures on *Naming and Necessity*: 'One should even remind oneself that the 'worlds' terminology can often be replaced by modal talk—"It is possible that..."' [16, p. 15]. Of course, Kripke is doing this remark in a specific context (where he explains rigid

---

[22] In fact, if you believe in a complete determination of the future, you will not have other possibilities than epistemic ones.

[23] This is a separate discussion, which, we believe, will help to resolve a lot of the philosophical problems concerning the understanding and status of 'laws'.

[24] Since we do not consider nested modalities, we do not need an accessibility relation.

designators), but the fact that he proposes the reading of possible worlds by use of plain possibility makes clear that they cannot serve, for Kripke, as *semantics* of possibility—otherwise one would have an obvious circularity.

6.5.5 A Cross Check

Let us have a look how we argue that something is *not* possible, i.e., how we argue for the sentence

(2) ϕ is not possible.

We think that such a sentence is not defended by reference to the fact that there could be no world in which ϕ holds—in contrast, we probably could easily imagine a world where ϕ holds.

Think of the example where somebody tells me 'Fiona is here in Coimbra.' However, since I just spoke with her on the phone when she was in Zürich, I will reply 'That is not possible.'

I do not exclude the possibility of a world in which Fiona is right now in Coimbra—even if she was 10 min ago in Zürich; but I reject the possibility on the grounds of my knowledge: Her presence in Coimbra would contradict my background theory, which includes some basic facts about travel times.

One might subsume this case under epistemic possibility, rephrasing my reply as 'I consider this as impossible.' And for epistemic possibility our objection to possible worlds semantics should be cogent: No one can seriously argue with different *worlds* to justify the possibility of the truth as well as the possibility of the falsity of Riemann's hypothesis.

But also when we argue for the impossibility of clearly non-epistemic statements we claim that we argue *primary* on the basis of contradictions:

It is impossible that bachelors are married.

This sentence is defended by the fact that a married bachelor contradicts its definition[25]—not because one could not think of a world in which a married bachelor exists.

6.6 An Alternative Analysis of Modalities: Necessity

Taking the standard duality of necessity and possibility, it would follow from our analysis of possibility, that a sentence ϕ is necessary, if (and only if) it is not independent. Since this includes the case that $\neg\phi$ is derivable, such an account is

[25] See the section on the normative character of unary necessity statements below.

obviously absurd. To keep the duality we would have to drop at least our pragmatic assumption that derivable sentences are not possible. Then, necessity would coincide with derivability. But this would turn necessity in a rather redundant notion.[26] Thus, the duality is in fact not the main issue.

6.6.1 Necessity as Binary Relation

We claim that necessity statements, as we find them in natural language use, are of a special form: They are relations between two sentences—or, at least, can be completed to such a relation. In other words, necessity is a binary relation on statements, which should be given in the form:

(3) ϕ is necessary for ψ.

As analysis for this sort of necessity statements we propose the following reading:

Every proof of ψ uses ϕ.

This reading of necessity is discussed in our paper [14]. It requires, of course, an explanation of the *use* of a formula in a proof. While its intuitive meaning seems to be clear, it turns out that a formal definition is far from being easy.[27] However, the notion of use is no problem if we restrict ϕ to be an axiom. In this case, ϕ is—somehow: has to be—used in every proof of ψ, if ψ is not provable in the axiom system from which ϕ is removed. Thus, our analysis seems to be restricted to sentences of the form (3), where ϕ is an axiom. But, we would like to put it in different terms: It is restricted to sentences of the form (3), for which ϕ *can be taken as an axiom*.[28] That means the analysis depends on the way we axiomatize the background.

6.6.2 Variety of Alternatives

If we consider all axioms, including the logical ones, we face a similar problem as modal logic with logical omniscience: The logical axioms might be necessary for

[26] In fact, at some occasions necessity is used in a redundant way with respect to the truth of a sentences (or argument). In these cases it should just serve as special emphasize. The most prominent example for such a use is probably the use of ἀνάγκη by Aristotle for the conclusion of syllogisms, cf. Łukasiewicz [18, §5].

[27] The problem starts already with the question what happens if we replace ϕ by a logically equivalent formula ϕ'; it might be that in this way we get a proof of ψ which does not use ϕ literally, but somehow implicitly.

[28] See the example in the following subsection.

every other statement. This seems to be fine for a notion of *logical* necessity; but it is normally inadequate for the use of necessity in natural languages.[29]

But not only the logical axioms are usually excluded from the discourse; we claim that for the analysis of necessity statements we make implicitly compromises which are the axioms considerable as necessary, depending on the context. Technically they can be grouped in a *variety of alternatives*, cf. [14, Def. 2].

To illustrate the variety of alternatives let us think of an example from a soccer league:

($\star$) A must win today to win the league.

We will dispense with the background theory which contains, on the one hand, the rules of building a soccer league from the result of the matches and, on the other hand, the results of the matches played so far. But let us assume that all other matches are already played and A is—one game behind—2 points behind the current leader B. Thus, with the three points of a win in the last match, it would go ahead and win the league. Now, the variety of alternatives will consist of the possible results of the last match in the form:

$$\{A \text{ wins; } A \text{ draws; } A \text{ looses}\}$$

From these three alternatives only 'A wins' will ensure the win of the league. Thus, ($\star$) turns out to be true.

One could object that the 'axioms' should not be 'A wins', 'A draws' and 'A looses' but the exact scores of the match like 'A wins 1:0', 'A plays 1:1', etc.; and the former formulas would be only derived. However, since the exact score is irrelevant for the question whether A can win the league, we can just take the plain distinction of win/draw/defeat as axioms; would the score be relevant, we would have to choose a more fine-grained variety. Thus, what actually counts as an *axiom* depends on the context.

Back to the example: it is easy to think of an alternative way in which A could win the league, even when loosing the last match: The soccer federation could ban B from the competition because of influencing referees. This is definitely a possible scenario in which A would win the league. But, under normal circumstances,[30] no soccer supporter would allow it as a serious argument for the falsity of ($\star$). Thus, the context of a (binary) necessity statement determines the variety of alternatives; other 'possible' alternatives are excluded from the discourse. In other words, the variety of alternatives is stipulated.

[29] However, there are limit cases: The question whether the *tertium-non-datur* is necessary for the proof of a certain mathematical theorem, is an example where we like to consider a logical axiom in this context. And, in fact, it is a prime example for our reading of necessity.

[30] We would have non-normal circumstances if the soccer federation already announced an investigation of the doings of B.

If we now compare our account with possible worlds semantics, one can easily realize that the variety of alternatives gives rise to a set of possible worlds.[31] However, as for the analysis of possibility, this semantic interpretation is *secondary* in comparison with the analysis of necessity in terms of proofs. In this view, again, possible worlds are just an epiphenomenon. And, instead of a quantification over possible worlds, we have a quantification over proofs; these proofs are well defined objects and do not need any ontological additions for our analysis.[32]

6.6.3 Unary Necessity

Now, we consider the use of necessity as a unary relation:

(4) ϕ is necessary.

We think that the majority of such statements would have to be extended to binary statements (if they should not just be considered as immediately false). For instance, when a politician says 'We must raise the taxes'[33] nobody can seriously consider such a statement as literally true.[34] But it is also not taken as a senseless sentences. We would propose that it has to be understood as something like 'We must raise the taxes to keep the budget under control while we keep our plans of expenses.' In such a reading the necessity statement might be true. In fact, this reading makes clear that the politician excludes the obvious alternative 'We could reduce the expenses' from the variety of alternatives.[35]

However, there are unary necessity statements which does not seem to be shortened forms of binary ones. One example could be '2+2 is necessarily 4'. However, it is an independent philosophical question in which way this statement is different from the plain fact that '2+2 is 4'. We think that the necessity here is either just an emphasis, or it is a generic attribute for mathematical (and logical) truth to indicate

---

[31] Thus, if you still prefer possible worlds semantics over our reading, you should at least consider to use the varieties of alternatives to determine your variety of possible worlds.

[32] If you object to proofs as formal objects in this analysis, you can replace it by the (less well defined) notion of *argument*; in particular for a notion of *epistemic necessity*, constructed in analogy to the epistemic possibility above, arguments—known to the defender of a necessity statements—could be more adequate than proofs.

[33] This is a more colloquial form of 'To raise the taxes is necessary' which would be (very close to) an instance of (3).

[34] In particular, in a possible worlds semantics, every tax payer could without any doubt imagine a world where the taxes do not raise.

[35] To go even a step further, we could suggest that the communicative content of a necessity statement is often just the exclusion of certain—perfectly reasonable—alternatives. In terms of possible worlds semantics: A necessity statement is not to be analyzed in terms of possible worlds; in contrast, it is meant to define the possible worlds which the utterer considers as 'possible'. This latter 'possible' might be better called 'preferable' or the like.

them as such.[36] We doubt that it should be analyzed in terms of possible worlds; in fact, we can easily think of a world in which 2+2 equals 1—just consider the group $\mathbb{Z}/3\mathbb{Z}$.[37] Most likely, the philosophical reply would be that with the switch to this other group the meaning of the symbols was changed, but for 'the real 2' the necessity statement would still hold. But this would force one to combine possible worlds semantics with a theory of rigid designators, a theory which makes the full approach even more dubious; non-rigid designators are precisely the point in possible worlds semantics to analyze the classical Phosphorus/Hesperus example.

6.6.4 The Normative Nature of Unary Necessity

Another classical example for unary necessity is 'It is necessarily the case that a bachelor is unmarried'. Again, taking a possible worlds semantics for it, we would consider it as plainly wrong, since we could easily imagine a world in which bachelors are the married men. To consider it as true, would turn this statement in a normative one: Worlds in which bachelors are married are excluded from the discourse. But how can we treat this kind of statements in our analysis? Since they are not binary, they are outside of the scope of the reading given above. However, taking up the normative character of these statements we can consider them as limit cases of our analysis in which the necessity expresses the *axiomatic* status we attribute to the statement. That means, with a statements of the form (4) we want to express that ϕ has the status of an axiom; it is put at the beginning of our discourse and cannot be questioned in the further discussion. Such a reading obviously fits well in this example if we think of the *definition* of bachelor as a men which is not married: definitions should have the status of axioms.[38]

Reflecting on this last reading—the normative character of unary necessity statements which should get an axiomatic status—we can even subsume the example of '2+2 is necessarily 4' under it: It excluded other readings of 2 (and + and 4) from the discourse. Even more, the example of raising taxes can also be subsumed: The

---

[36] This is in line with Frege when he writes: 'Das apodiktische Urtheil unterscheidet sich vom assertorischen dadurch, dass das Bestehen allgemeiner Urtheile angedeutet wird, aus denen der Satz geschlossen werden kann, während bei den assertorischen eine solche Andeutung fehlt. Wenn ich einen Satz als nothwendig bezeichne, so gebe ich dadurch einen Wink über meine Urtheilsgründe.' [10, p. 4] ('The apodictic judgment differs from the assertory in that it suggests the existence of universal judgments from which the proposition can be inferred, while in the case of the assertory one such a suggestion is lacking. By saying that a proposition is necessary I give a hint about the grounds for my judgment.' [11, p. 13].)

[37] If one allows the symbol 4 in this group, of course, 4 equals 1 and one could try to save this particular necessity statement; but it is clear that a statement like '2+2 is necessarily different from 1' would break down.

[38] Obviously, this view is closely related to Carnap's *meaning postulates*.

politician excludes the question of raising taxes from the discourse and gives the fact that the taxes will raise a normative status.[39]

Also this analysis of unary necessity statements is compatible with possible worlds semantics; but it not only has the advantage of being ontologically modest, it even led us to the insight that necessity statements might have to be considered as normative.

6.7 The Temporal Aspect

6.7.1 The Dynamics of the Axiom System

In connection with our pragmatic principle for possibility we already addressed a temporal aspect: When we have a statement about a point t in the future, its possibility is no longer an issue when the time arrives at point t; it is decided and it can be checked whether it is true or false, but possibility is no longer of interest.[40]

Now, we look also to a temporal aspect in connection with necessity statements. Let us consider the statement:

(5) ϕ is no longer necessary.

We should expect that a semantics of necessity could offer something to its understanding. Possible worlds semantics has not much to offer for it. It is clear that ϕ has to be a sentence which was holding in all possible worlds at some point in the past, but now there is at least one world in which does not hold any longer. Thus, we would need a *dynamics* in the worlds which are under consideration. We already criticized above that possible worlds semantics suffers from the missing specification of the variety of worlds. But if we have no clear specification of the variety, we have even less hope to get a controlled dynamics for it.

In our reading, such a dynamics is easy to implement. Let us first look to the binary companion of (5), 'ϕ is no longer necessary for ψ.' In this case ϕ has to be an axiom taken from a variety of alternatives. Now, if we allow the axiom system to change, for instance by adding new axioms (which could express new fact about the world coming true in the meanwhile), it is easy to imagine that we can get a new proof of ψ which could use one of the new axioms χ, but not needing ϕ any longer. The point here is that the change of the axioms takes place in a very controlled way. It should even be possible to say *why* ϕ is no longer necessary for ψ: because, now we have χ.

For the case of unary necessity the situation is not much different; there might be new axioms which make the normative character of ϕ needless. Assuming an unex-

[39] It would be probably more honest if the politician would say 'We *will* raise the taxes' instead of 'We must raise the taxes'; but it is obvious that it is politically smarter to use the second statement.

[40] This does not hold for epistemic possibility for the case we *did not* check the truth or falsity.

pected growing of the economy, our politician would probably be more than happy to announce: ‘It is no longer necessary to raise the taxes’. In the new (axiomatic) situation, he considers his normative character of the tax-raising as obsolete.[41]

6.7.2 Nested Modalities

So far, we restricted ourselves to sentences which do not contain nested modalities. Even if we claim that most of the examples of nested modalities, discussed in the context of modal logic and possible worlds semantics, are rather artificial, there are surely exceptions. Let us just look at the following simple example:

(6) It is possible that ϕ is necessary.

We think that there are two types of this statement. On the one hand, it might involve a temporal aspect: At the moment, ϕ is not necessary, but later on—with new ‘axioms’ around—it can become necessary. This is somehow the inverse of the case described in the previous subsection. On the other hand, it can be considered as possibility on the (next) meta-level: The possibility in the first part of the sentences refers to the fact that *it is possible that there is a background theory in which ϕ turns out to be necessary*. Thus, we have to take different background theories into account—just as we do when deal with the temporal aspect.[42]

6.8 Conclusion

In this paper we criticized the classical approach of modal logic and possible worlds semantics to the modalities *possible* and *necessary*. The main criticism is that this approach does not provide us with a satisfactory theory to analyze the use of these modalities in natural language; it rather allows only to study some algebraic relations of nested modalities. This criticism is also given in a more provocative way in our contribution to the Festschrift for Shahid Rahman [15]; for other critical accounts to possible worlds see, for instance, [9, 28].

As alternative we outlined a reading of possibility and necessity as meta-statements about an axiom system of the current situation. This view does not only

[41] Such a dynamics should, of course, not apply to the logical and mathematical examples; however, it might apply to the ‘definitional’ examples: ‘civil marriage is a contract between a man and a woman’ might have been considered as a necessary truth due to the definition of marriage. Today, several countries have legislation redefining this notion.

[42] Defenders of possible worlds semantics may point out that different forms of modalities require also different varieties of possible worlds. We are not aware of any approach which combines such different modalities in a single framework; in fact, such a framework would probably have quite drastic implications for the underlying realism.

overcome several of the notorious problems of modal logic and possible worlds semantics, but it allows also to describe some aspects of modalities more adequately. As a particular advantage we consider the fact that *en passant*, modalities are integrated in the theory of contingent facts. It should be clear that the given analysis is quite promising to resolve also other, more problematic intensional phenomena, as, for example, *counterfactuals*. This, however, is still work to do.

There are at least two other 'proof-theoretic' approaches to modalities in the literature which are worth studying in relation to the account presented here. One is Artemov's *logic of proofs*, recently put forward explicitly as a *justification logic*, [1, 2]. The other is the interpretation of modalities in *contextual judgment calculi* given by Davies, Nanevsky, Pfenning, and Pientka, [20, 21]. In the work of Primiero [22, 23] one can even find some kind of combination of these two approaches. All of them put forward an epistemic reading of modalities similar to the one given in this paper, but they internalize the modalities—as boxes or other constructions—in the object language.

Let us finish with the observation that our approach dispense with the realism coming usually with possible worlds semantics. Since our analysis lifts modality statements on the same level as provability, it does not require any 'new' philosophical ontology: The approach will use exactly the same ontology one uses to interpret provability. Thus, it shows that, modulo provability, modalities are neutral with respect to realism/antirealism.

Acknowledgements I thank Jesse Alama, Volker Halbach, Greg Wheeler, Bartosz Więckowski, and John Woods for helpful comments on earlier versions of this paper. Research partially supported by the ESF research project *Dialogical Foundations of Semantics* within the ESF Eurocores program *LogICCC* (funded by the Portuguese Science Foundation, FCT LogICCC/0001/2007).

References

1. Artemov, S., and R. Iemhoff. 2007. "The Basic Intuitionistic Logic of Proofs." *Journal of Symbolic Logic* 72(2):439–51.
2. Artemov, S., and E. Nogina. 2005. "Introducing Justification to Epistemic Logic." *Journal of Logic and Computation* 15(6):1059–73.
3. Bull, R., and K. Segerberg. 1984. "Basic Modal Logic." In *Handbook of Philosophical Logic*, edited by D. Gabbay and F. Guenthner, vol. II, 1–88. Dordrecht: Kluwer.
4. Carnap, R. 1947. *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago, IL: University of Chicago Press.
5. Chihara, C. S. 1998. *The Worlds of Possibility*. Oxford: Oxford Unviersity Press.
6. Copeland, B. J. 2002. "The Genesis of Possible Worlds Semantics." *Journal of Philosophical Logic* 31(2):99–137.
7. Divers, J. 2002. *Possible Worlds*. London: Routledge.
8. Føllesdal, D. 1980. "Eintrag *Semantik*." In *Handbuch wissenschaftstheoretischer Begriffe*, edited by J. Speck, vol. 3 (R–Z), 568–79. Göttingen: Vandenhoek and Ruprecht.
9. Forster, T. 2005. "The Modal Aether." In *Intensionality,* vol. 22 of *Lecture Notes in Logic*, edited by R. Kahle, 1–19. Wellesley, MA: ASL and AK Peters.
10. Frege, G. 1879. *Begriffsschrift. Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Louis Nebert.

11. Frege, G. 1879. *Begriffsschrift*, a Formula Language, Modeled Upon That of Arithmetic, for Pure Thought, 1–82. English translation of *Begriffsschrift. Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Louis Nebert; reprinted in *From Frege to Gödel a Source Book in Mathematical Logic, 1879–1931*, edited by Jean van Heijenoort. Cambridge, MA: Harvard University Press.
12. Grice, P. 1975. "Logic and Conversation." In *The Logic of Grammar*, edited by D. Davidson and G. Harman, 64–75. Encino, CA: Dickenson.
13. Kahle, R. 2002. "Structured Belief Bases." *Logic and Logical Philosophy* 10:49–62.
14. Kahle, R. 2006. "A Proof-Theoretic View of Necessity." *Synthese* 148(3):659–73.
15. Kahle, R. 2008. "Against Possible Worlds." In *Dialogues, Logics and Other Strange Things. Essays in Honour of Shahid Rahman*, vol. 7 of *Tributes*, edited by C. Degremont, L. Keiff, and H. Rückert, 235–53. London: College Publications.
16. Kripke, S. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
17. Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
18. Łukasiewicz, J. 1957. *Aristotle's Syllogistic. From the Standpoint of Modern Formal Logic*, 2nd Edition. Oxford: Clarendon Press. (1st Edition, 1951).
19. Mares, E. D. 2004. *Relevant Logic*. Cambridge, MA: Cambridge University Press.
20. Nanevski, A., F. Pfenning, and B. Pientka. 2008. "Contextual Modal Type Theory." *ACM Transactions on Computational Logic* 9(3):1–49.
21. Pfenning F., and R. Davies. 2001. "A Judgmental Reconstruction of Modal Logic." *Mathematical Structures in Computer Science* 11:511–40.
22. Primiero, G. 2009. "Epistemic Modalities." In *Acts of Knowledge—History, Philosophy and Logic*, edited by G. Primiero and S. Rahman, 207–32. London: College Publications.
23. Primiero, G. 2010. "Constructive Contextual Modal Judgments for Reasoning from Open Assumptions." In *Programs, Proofs, Processes*, edited by F. Ferreira, H. Guerra, E. Mayordomo, and J. Rasga. *Proceedings of the Sixth Conference on Computability in Europe, CiE 2010*, 336–45. University of Azores, Centre for Applied Mathematics and Information Technology.
24. Smith, R. 2009. "Aristotle's Logic." In *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition), edited by Edward N. Zalta. http://plato.stanford.edu/archives/sum2011/entries/aristotle-logic/
25. Stalnaker, R. 1995. "Modalities and Possible Worlds." In *A Companion to Metaphysics*, edited by J. Kim and E. Sosa, 333–37. Oxford: Blackwell.
26. Wehmeier, K. 2004. "In the Mood." *Journal of Philosophical Logic* 33:607–30.
27. Wehmeier, K. 2005. "Modality, Mood, and Descriptions." In *Intensionality*, vol. 22 of *Lecture Notes in Logic*, edited by R. Kahle, 187–216. Wellesley, MA: ASL and AK Peters.
28. Woods, J. (preprint). "Making Too Much of Possible Worlds." 201x. Available at the author's home page: http://www.johnwoods.ca/

Chapter 7
Antirealism, Meaning and Truth-Conditional Semantics

Neil Kennedy

7.1 Introduction

The notion of meaning is the central concern of Dummett's philosophy. It is through this notion that he develops his critical stance on realism, and through this notion that he argues, notoriously, for the rejection of classical logic. From Dummett's criticism of realism emerges a position he coins "antirealism", a thesis opposed to realism and slightly different from the traditional opponent to realism, which is commonly known as "idealism". In this paper, I would like to re-examine Dummett's arguments against realism and, in particular, his arguments against truth conditional semantics. Dummett claims that a (realist) truth conditional meaning theory will invariably encounter limitations when accounting for the meanings of the statements of the so-called *disputed class*. It is by reflecting on the statements of this class that Dummett arrives at the various canons of his philosophy, which are: the rejection of truth conditions, holism, and classical logic on the one hand, and the adoption of verificationism, molecularism and intuitionistic logic on the other. The first task of this paper, then, will be to faithfully reconstruct Dummett's arguments on meaning. The second task will be critical. I will attempt to expose the shortcomings of an antirealist theory of meaning, and show how truth conditional semantics (be they realist or not) emerge relatively unscathed from Dummett's criticism.

7.2 Dummett's Antirealism

In Dummett's view, the debate between realism and antirealism is not on the kind of entities that exist and the properties they exemplify (and, if I may add, *how* they exemplify them), that is, a debate concerning ontological matters proper, it is a debate concerning the meanings of statements and the relation of meaning to truth and verification (or proof). The idea that (the meanings of) statements are more basic

N. Kennedy (✉)
Université du Québec à Montréal, Montréal, QC, Canada

Université de Paris 1–IHPST, Paris, France
e-mail: neil.patrick.kennedy@gmail.com

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_7,

than objects is a consequence of the espousal of Frege's dictum (or context principle) according to which "only in the context of a sentence does a name stand for anything".[1] Thus realists and antirealists do not (or should not) concern themselves with a class of disputed *entities* but rather with a class of disputed *statements*, and what meanings the statements belonging to this class should have. In this debate, realism is characterised as

> the belief that statements of the disputed class possess an objective truth-value, independently of our means of knowing it: they are true or false in virtue of a reality existing independently of us. [...] That is, the realist holds that the meanings of statements of the disputed class are not directly tied to the kind for them that we can have, but consist in the manner of their determination of as true or false by states of affairs whose existence is not dependant on our possession of evidence for them.[2]

This is opposed to the antirealist who claims

> that the meanings of these statements are tied directly to what we count as evidence for them, in such a way that a statement of the disputed class, if true at all, can be true in virtue of something of which we could know and which we should count as evidence for its truth.[3]

So,

> [t]he dispute concerns the notion of truth appropriate for statements of the disputed class; and this means that it is a dispute concerning the kind of *meaning* which these statements have.[4]

It is thus by meaning and meaning alone that this opposition can be resolved: in the right corner, the realist defends a notion of meaning based on truth or truth conditions and does not preoccupy himself with verification or proof, and in the left-corner, the antirealist insists that a theory of meaning is incorrect unless meaning is construed in terms of possession of verification or proof.

Why does the Dummettian antirealist claim that meaning must involve verification? To understand the relation between meaning and verification, we must take a moment to consider the thesis according to which meaning is use. In the *Philosophical Investigations*, Wittgenstein states that "[f]or a *large* class of cases—though not for all—in which we employ the word 'meaning' it can be defined thus: the meaning of a word is its use in the language".[5] The meaning of a word, expression or statement is thus constituted (for a large class of cases) by the way we use this word, expression or statement in our language activities. Dummett applies this to the meaning of mathematical statements:

---

[1] M. Dummett, *The Philosophical Basis of Intuitionistic Logic*, [5, p. 230]. Henceforth, *The Philosophical Basis of Intuitionistic Logic* will go by the name TPBOIL. Many of the arguments of TPBOIL can be found developed in a similar fashion in *The Logical Basis of Metaphysics* [6].

[2] M. Dummett, *Realism*, [5, p. 146].

[3] M. Dummett, *Realism*, [5, p. 146].

[4] M. Dummett, *Realism*, [5, p. 146].

[5] [19, p. 43].

> The meaning of a mathematical statement determines and is exhaustively determined by its *use*. The meaning of such a statement cannot be, or contain as an ingredient, anything which is not manifest in the use made of it, lying solely in the mind of the individual who apprehends that meaning: if two individuals agree completely about the use to be made of the statement, then they agree about its meaning. The reason is that the meaning of a statement consists solely in its role as an instrument of communication between individuals, just as the powers of a chess-piece consist solely in its role in the game according to the rules.[6]

We can read this as a reduction of meaning to the observable behaviour of language speakers. Dummett is quite clear on the point that nothing in meaning can transcend this observable use of language, that no innate ingredient, for example, could contribute to our understanding of language, for it would result in failure of communication.[7]

It is then stated that a "model of meaning is a model of understanding, i.e. a representation of what is known when an individual knows meaning".[8] There is a certain ambiguity in this assertion. What Dummett is saying is that a theory of meaning must specify, for every statement *S* of a given language, the knowledge (implicit or explicit) an agent has when he is *said* to know the meaning of *S*. Combined with the meaning is use principle, this boils down to the idea that "[t]o grasp the meaning of an expression is to understand its role in the language".[9]

The next step in the argument against the realist position is the claim that it is violating a fundamental aspect of a theory of meaning, and so the realist theory of meaning will eventually fail. For a mathematical sentence outside the disputed class, hence a decidable sentence, there is no problem because

> a grasp of the condition under which the sentence is true may be said to be manifested by a mastery of the decision procedure, for the individual may, by that means, get himself in a position in which he can recognise that the condition for the truth of the sentence obtains or does not obtain [. . .].[10]

However, when a sentence is undecidable (i.e. belongs to the disputed class), "as is the case with the vast majority of sentences of any interesting mathematical theory",[11] then

> the condition which must, in general, obtain for it to be true is not one which we are capable of recognising whenever it obtains, or of getting ourselves in a position to do so.[12]

---

6 *TPBOIL*, [5, p. 216].

7 It could be objected here that Dummett is unjustifiably eliminating the possibility that our mental and cerebral make-up contributes in important ways to language use and, consequently, to meaning. Environmental factors only do not explain why the cat, dog or monkey do not communicate the way we humans do. If Dummett concedes this, as he must, why claim then that language understanding "consists solely in its role as an instrument of communication"?

8 *TPBOIL*, [5, p. 217].

9 *What is a Theory of Meaning? (I)*, [7, p. 2]. Henceforth *WIATOM (I)*.

10 *TPBOIL*, [5, p. 225].

11 *TPBOIL*, [5, p. 225].

12 *TPBOIL*, [5, p. 225].

The realist's meaning therefore runs into trouble because

> if the knowledge that constitutes a grasp of the meaning of a sentence has to be capable of being manifested in actual linguistic practice, it is quite obscure in what the knowledge of the condition under which a sentence is true can consist, when that condition is not one which is always capable of being recognised as obtaining.[13]

His truth-conditional theory of meaning can only accommodate "those statements which are in principle effectively decidable".[14]

Let us briefly go over the steps in this argument again. First, there is the claim that a theory of meaning is a theory of understanding: meaning is whatever is grasped or known by the speaker when he correctly uses language. Second, there is the realist claim according to which the meaning of a statement is given by the truth conditions of the statement, conditions that need not involve any epistemic constraints whatsoever. Third, there is the claim that certain statements, statements of the disputed class, are such that their (realist) truth conditions can obtain without us being capable of knowing that they obtain (either for practical or theoretical reasons). But these three claims are jointly inconsistent: if meaning is construed as truth conditions, and if the obtaining of these truth conditions is unknowable, then realist meaning can not be manifested in language practice, contrary to the first claim. Dummett thus feels justified in rejecting the second, realist claim to restore consistency.

Moving along, Dummett proposes to replace the notion of truth by that of proof as that central feature of use on which our theory of meaning is to be built: "a grasp in the meaning of a statement consists in a capacity to recognise a proof of it when one is presented to us".[15] This is stated elsewhere in a stronger form as the fact that "we know the meaning of a mathematical statement if and only if we know what to count as a proof of it".[16] The move from truth to proof is closely followed by the move from truth conditions to proof or assertability conditions: the meaning of a statement is no longer explained by the conditions that must obtain for this statement to be true but "by stipulating when it [the statement] may be asserted in terms of conditions under which its constituents may be asserted".[17] Proof is the central feature that governs the use of assertions: in order to make an assertion, one must be able to produce a proof of it; in order to understand an assertion, one must able to recognize a proof of it. So meaning is proof and the grasping of meaning is the recognition of proof.

These remarks on meaning are not limited in scope to mathematical statements for they could "just as well have been applied to any statements whatever",[18] "the

---

[13] *TPBOIL*, [5, p. 224].

[14] M. Dummett, *Truth*, [5, p. 24].

[15] *TPBOIL*, [5, p. 225].

[16] M. Dummett, *Realism*, [5, p. 153].

[17] *Truth*, [5, p. 18].

[18] *TPBOIL*, [5, p. 226].

appropriate generalisation [...] would be the replacement of the notion of truth [...] by that of verification".[19] So a theory of meaning of the kind described above is, according to Dummett, quite general and can accommodate all types of statements.

Why is a theory of meaning based on assertability conditions better than one based on truth conditions? We saw that a realist theory of meaning encounters problems when it comes to accounting for the meanings of disputed class statements. Dummett insists repeatedly on the fact that "a realist interpretation is possible only for those statements which are in principle effectively decidable",[20] we thus expect, by contrast, that an *antirealist* interpretation is possible for undecidable statements. How is this accomplished? If I understand Dummett correctly, here is where harmony steps in.

Harmony is discussed in relation to holism, a thesis Dummett considers false. His main argument for rejecting it is once again taken from observations on the transmission of language: if meaning were holistic, than nothing less than the entire use of language would be necessary in order to grasp the meaning of a term or a statement. This is because no individual meaning is available for statements in a holistic framework, and if no individual meaning is available then how is one to progressively grasp language use? Tennant [17, chap. 4] characterises the situation as follows: both Quine and Dummett agree on the fact that meaning is not a transcendental entity, but they diverge on the question of the individuation of meaning. For the first, statements do not have determinate meaning, they only have a position in the web, which is the only true meaning-bearing entity for a holist; whereas for the latter, statements have determinate and individual meaning.

How does Dummett bypass the conclusions the holist arrives at, namely the idea that individual meaning could only be given to simple observation sentences? This can be accomplished, it is claimed, by requiring that "language as a whole be a conservative extension of that fragment of the language containing observation statements".[21]

To understand this, we must return to the two aspects deemed essential in the use of sentences: introduction and elimination rules for the logical connectives. The term harmony—coined by Dummett himself—more or less refers to the equilibrium between introduction and elimination rules of a logical constant, and more generally, in the context of philosophy of language, to the equilibrium between the different aspects of use. What Dummett precisely understands by harmony was, and is to this date, subject to much debate.[22] Suffice it to say that he considers it to be intimately related, if not identical, to the notion of conservative extension, and

---

[19] *TPBOIL*, [5, p. 227].

[20] *Truth*, [5, p. 24].

[21] *TPBOIL*, [5, p. 221].

[22] Tennant [17, chap. 10] surveys the various interpretations the terms harmony and conservative extension have been given by Burgess, Grandy and Prawitz.

especially the way it was put to use in Belnap's response to Prior's pathological connective (cf. [14]):

> We may now state the demand for the consistency of the definition of the new connective, *plonk*, as follows: the extension must be *conservative*; i.e. although the extension may well have new deducibility-statements, these new statements will all involve *plonk*.[23]

If γ is some new connective we add to a language L, the elimination rule of γ is in harmony with its introduction rule iff, for all formulae A in which γ does not occur,[24]

If $\vdash_{L(\gamma)} A$, then $\vdash_{L} A$.

This condition assures us that nothing new can be proved in the old fragment with the new connective. Phrased in the language of semantics: the meaning of γ is given exclusively by its introduction and elimination rules, it does not 'contaminate' the meanings of the other constants.

The fact that the int-elim rules of a logical connective are in harmony allows us to spell out Dummett's response to holism, which is known as molecularism.[25] On a first approximation, molecularism is the thesis according to which each statement can be associated to a unique and determinate meaning: it is supposed first that basic statements admit basic meanings, then, from these meanings and the way harmonious connectives operate, a determinate meaning can also be given to complex statements. The definite proposal is something along the lines of BHK semantics amended with the notion of canonical proof.

A canonical proof in natural deduction is a normalized proof, a proof in which detours have been eliminated. The traditional BHK semantics takes a proof of $A \vee B$, say, to be a proof of A or a proof of B. It is of course *sufficient* to have a proof of A or a proof of B in order to have a proof of $A \vee B$. But if we want this definition to have any scope, we will require that in order to have a proof $A \vee B$ we have a proof of A or a proof of B, i.e. that a proof of $A \vee B$ *necessarily* implies a proof of A or a proof of B. In the case of intuitionistic logic at least, normalisation comes to our rescue, for it enables us to make a canonical proof of $A \vee B$ out of any proof of $A \vee B$, and therefore we obtain a (canonical) proof of A or a (canonical) proof of B. This is where harmony closes in on classical logic: only a *harmonious* logic will allow us to define a notion of canonical proof and classical logic is not harmonious (or so it would appear).

---

23 [1, p. 131].

24 To be correct, one would have to insist on the difference between language and theory in this definition. The notion of conservative extension involves both: the language $L(\gamma)$ is an extension of the language L, and the theory $T(\gamma)$ associated with $L(\gamma)$ is also an extension of the theory T associated with L. And note that many theories can be associated to the same language.

25 For more details on int-elim rules, see [12].

Dummett and Prawitz's upgraded version of BHK semantics goes like follows[26]:

REVISED 'CANONICAL' BHK SEMANTICS

To form a canonical proof of	It is necessary and sufficient to have
$A \wedge B$	A pair (σ, τ) where σ and τ are canonical proofs of A and B resp.
$A \vee B$	A canonical proof σ of A or a canonical proof τ of B.
$A \rightarrow B$	A procedure which applied to a canonical proof σ of A yields a canonical proof τ of B.
$\exists x A$	A pair (τ, σ) where τ is a term and σ a canonical proof of $A[x/t]$.
$\forall x A$	A procedure which applied to a term t yields a canonical proof σ of $A[x/t]$.

It is important to realise that Dummett's rejection of classical logic is the last step of the foregoing arguments concerning meaning. First, there was the rejection of realist truth conditions, then came the idea that grasp of meaning is constituted by the recognition of proof or verification, and finally came the idea of harmony and that of meaning via canonical proof. In order to have canonical proofs (normalisable proofs), the logical constants must be harmonious and *this* is where classical logic fails. Dummett's rejection of classical logic is thus a very special corollary of his positions on meaning.

7.3 Harmony and Classical Logic

Harmony can be likened to a whole constellation of proof-theoretical notions besides that of conservative extension. It can be understood as the principle of inversion, separation, normalisation, or the *Hauptsatz*, the latter being the most general version of the idea. As Gentzen observes,

> [t]he *Hauptsatz* says that every purely logical proof can be reduced to a definite, though not unique, normal form. Perhaps we may express the essential properties of such a normal proof by saying: it is not roundabout. No concepts enter into the proof other than those contained in its final result, there use was therefore essential to the achievement of that result.[27]

It is this feature that Dummett seeks for a molecular conception of language: separability of meaning. As was shown above, the revision of classical logic in favour

[26] As usual, negation is defined with $\perp$ and $\rightarrow$.
[27] [10, p. 69].

of intuitionistic logic comes from the fact that logical constants of the latter are separable whereas classical ones are not: proofs in NJ are normalisable but proofs in NK aren't. One can also see this in the conservativeness angle of separability: the addition of classical negation rules to the fragment containing only the constants '$\wedge$', '$\vee$', '$\rightarrow$', '$\forall$', '$\exists$' (which are defined in the same fashion in NJ and NK) does not produce a conservative extension. For example, Pierce's law, i.e. $((P \rightarrow Q) \rightarrow P) \rightarrow P$, is provable in the extension but not in the negation-less fragment (and there are instances of both that don't contain the negation sign).

As we have seen, Dummett isolates two aspects of use that he considers of importance for a correct rendering of the semantics of (mathematical) statements. And he believes that the requirement of harmony between these two aspects of use is strong enough of a constraint to fix a definite choice of logic, which as it turns out is intuitionistic logic. Let us first see how this constraint of harmony falls short of its promises.

Why don't the proof-theoretical constraints suffice to establish the desired conclusion? First of all, and this is a minor observation, rejection of classical logic need not mean acceptance of intuitionistic logic. In this respect the argument is underdetermined, for intuitionistic logic is not the only harmonious logic. In fact, amongst harmonious logics, one could say that minimal logic is an even better choice, because proofs in this system are not only normalisable, they are uniquely so. But this is no major problem, for as long as classical logic is not harmonious, one can still defend rejection of classical logic in favour of another (weaker) logic, although further criteria are necessary in other to fix the choice.

Tennant [17], for instance, pushes for the adoption of relevant intuitionistic logic, arguing that it makes a Popperian conception of science possible. Prawitz, on the other, does not push for logical revision in a way that is as strong as this. He only points to the fact that, in being harmonious, intuitionistic logic is superior to classical logic, and in the end, after some process of reflexive equilibrium, the first should have better chances of winning over the second [13, p. 17].

What truly jeopardizes the Dummett-Prawitz argument is its apparent dependence on the specific presentation one gives to a language (and its underlying logic). Take the following simple example: suppose that instead of defining classical logic with the usual conjunction, disjunction, implication and negation constants (and bottom), we use Sheffer's stroke (and bottom). It is well known that classical propositional logic can be completely defined with this connective only (and one can even present it in natural deduction, cf. [16, 18]). In this presentation, classical logic trivially satisfies the conservative extension criterion for: (1) any complex formula A already contains the Sheffer stroke, and (2) no propositional variable is a theorem. Hence classical logic, in the Sheffer presentation, would be just as harmonious as intuitionistic logic.

One need not have recourse to Sheffer's stroke to illustrate this point. Gentzen's *Hauptsatz*, the counterpart to normalisation in sequent calculus, holds for intuitionistic *and* classical sequent calculus. The failure of normalisation—the natural deduction equivalent of the Hauptsatz—in the classical case is, once again, due to the specific nature of the presentation of classical natural deduction. Recall that LJ is the special case of LK where we restrict the consequent to at most one formula. It turns out that there is a way of presenting natural deduction in a fashion that reflects this specific relationship, one that is not at all apparent in the usual presentation of natural deduction. The idea, which is due to Boričić [3], is to have natural deduction operate on multisets of formulae instead of (single) formulae. New ‘structural rules’ must be added in order to deal with these multisets, which correspond to the weakening, permutation and contraction rules of sequent calculus, whereas introduction and elimination rules for logical constants stay more or less to the same. The result is a natural deduction system for classical logic NL in which proofs are normalisable.

One can prove Pierce’s first and second laws in the positive fragment (the fragment without negation) of NL, showing that NL is a conservative extension of its positive fragment[28]:

1	P	Assumption	
2	P, Q	1, Weakening	(1)
3	$P, P \to Q$	1-2, $\to$-intro	
4	$(P \to Q) \to P$	Assumption	
5	P, P	3,4, $\to$-elim	(4)
6	P	5, Contraction	(4)
7	$((P \to Q) \to P) \to P$	4,6, $\to$-intro	

Hence, according to the Dummett-Prawitz criterion, classical logic would be harmonious, that is, in no need of revision. Read was the first to observe the philosophical consequences these results have for Dummettian antirealism:

> [T]he formalism does not reveal the true nature of the disagreement between classicist and intuitionist, and in particular, harmony and autonomy are interesting properties open to both.[29]

And as he concludes further on:

> [T]he constructivist can still mount a challenge to classical logic. But we now see where that challenge should be concentrated—and where it is misguided. The proper challenge is to Bivalence, and to the classical willingness to assert disjunctions, neither of whose disjuncts is separately justified.[30]

A last observation on this matter. Unsurprisingly, normalisability is not the only property that depends on the specific presentation of a logical system. One also

28 The derivation in multiple formulae natural deduction is presented here in linear form but Boričić [3] presents it in the tree-like form more common to natural deduction. The tree can be readily extracted from the linear presentation below.

29 [15, p. 151].

30 [15, p. 151].

expects harmony, inversion and conservativeness to depend on such presentations. What is surprising, however, is that these properties don't always vary coextensively. For example, the invertible rules that one has in NJ do not translate into invertible rules in LJ. (A rule is invertible in LJ iff: if a deduction of a sequent with main connective γ can be given, then a deduction of the premises of the corresponding rule for γ can also be given.) For example, we have that $\rightarrow$-left is invertible in LK but not in LJ. (In NJ, one can only show a partial inversion for this connective.)

7.4 Antirealist Meaning and Holism

The second major item on the list of Dummett's premises is the claim that holism makes the learning of language impossible. I will argue here that a molecular theory of meaning will encounter a similar difficulty, though not at the same level. Granted, the meaning of a complex statement is given by the meaning of the composing sub-statements (and the meaning of the connectives), but once we reach the atomic level, the meanings of the sub-propositional expressions is left to Frege's context principle:

> a grasp of the meaning of any expression smaller than a sentence must consist in a knowledge of the way in which its presence in a sentence contributes to determining what is to count as a proof of that sentence.[31]

This is what makes Dummett's theory a molecular one rather than an atomic one.

The reliance on the context principle makes the meaning of terms and other sub-propositional expressions quite radically holistic, which is strange considering that the main impulse for the adoption of harmony was precisely the rejection of holistic meaning. If I were to exaggerate somewhat, I would even say that this term or sub-propositional holism is usually what meaning holism is all about, not holism about logical connectives. So whatever flaw one can find in the logical-connective holism of a classical theory, one can find it wholesale in the term holism advocated by Dummett's molecularism. The typical criticism that meaning holism receives from its opponents, a criticism parallel to Dummett's considerations on the learnability of language, is its inability to account for the productivity of meaning or thought.[32] Although separable logical constants, in molecular semantics, give us a hold on how the meaning of a complex statement is the product of the meaning of atomic ones, nothing that amounts to productivity can be said for sub-propositional expressions.

Maybe it would help the antirealist case if, instead of just focusing on proof recognition/production capacities, we were to consider the capacities involved in the recognition/production of terms, formulae, axioms and other linguistic expressions involved in making proofs. In this respect, the behavioural checkpoints of a

---

[31] *TPBOIL*, [5, pp. 225–26].

[32] Expression used in [9].

meaning theory *à la Prawitz* (cf. [13, p. 9]) could include, for example, the capacity to recognize/produce a term, as well as a capacity to recognize/produce a constant (symbol) and a function (symbol), etc. These capacities are far from being insignificant auxiliaries to the full use of natural deduction. If one were unable to distinguish a term from any other sub-propositional expression, he would be utterly incapable of applying the introduction and elimination rules for the universal and existential quantifiers. Furthermore, without a capacity to recognize predicates and atomic formulae, statements would go unnoticed and we would have nothing to apply our logic to.

We can give the benefit of the doubt to Dummett that knowledge of these term producing and recognition capacities was what he had in mind by the expression "knowledge of the way in which its [the term] presence in a sentence contributes to determining what is to count as a proof of that sentence". As far as I can see, this would be the closest one could get, in a Dummettian perspective, to a theory of individual meaning for terms (and other sub-propositional expressions). However, it is one thing to have term-formation and term-recognition capacities, but another to have a "grasp of the meaning of any expression smaller than a sentence". It seems clear to me that one can possess the necessary syntactic know-how to recognise that "horse" behaves like a common count-noun without knowing what "horse" refers to or means.

Take the term "0". Term formation rules for numerals will dictate what expressions are to be considered as numerals. For the rest, we must let Frege's dictum speak. In Dummett's idiom, the meaning of "0" is the way its presence contributes to the proof of a statement in which it occurs. Axiomatisations of arithmetic usually include the formulae "$\forall x(x+1 \neq 0)$" and "$\forall x(x \cdot 0 = 0)$", which incidentally involve the term "0". So in what way does '0' contribute to a proof of these axioms? The answer is: in no specific way whatsoever. Axioms are proofs of themselves, or they are considered unprovable assumptions. We either end up with a circular contribution to the proof (if the axiom is a proof of itself, the only way "0" contributes to the proof is by occupying the same syntactic position it occupies in the axiom), or we end up with no contribution at all. (Perhaps it is possible to do arithmetic without axioms, solely with introduction and elimination rules, but Dummett does not provide us with a set of rules that would accomplish this.)

Furthermore, proofs of statements that are logical validities will have nothing to do with the terms that may occur in them. This is a fundamental feature of a logical notion: logicality owes nothing to any peculiar nature or interpretation, the truth of a logical validity is preserved when we substitute (uniformly) the terms that occur in it for other terms. Here again, a term does not contribute to the proof of the statements in which it occurs. The molecular meaning of terms is thus very thin if anything at all.

This should make us realize that the meaning of a basic statement, on the molecular view, is also quite thin. We were told by Dummett that meaning of a complex statement is built from the meanings of the component statements, but nothing has been said of the meanings of the basic statements. As we saw with the examples above, a basic statement does not have a proof, or if it has one, it is of a very trivial nature. This is the base clause a theory of meaning requires, and Dummett offers us no precisions on how it may go. In a digression on Davidson's truth theoretical semantics, Dummett argues that Davidson can provide at most a "*modest* theory of meaning", that is, a theory of meaning that does not explain the grasp the primitive expressions of the language (*WIATOM (I)*, [7, p. 5]). In a modest theory of meaning, knowledge of the meaning of primitive expressions (be they atomic predicates or terms) is presupposed in the meta-language (or deferred to another theory), and so such a theory fails to give an exhaustive account of language use. A *full-blooded* theory of meaning, however, provides this sort of account, and it is a full-blooded theory of meaning that we must seek to provide. It seems, from our observation on basic statements, that Dummett fairs no better than Davidson on this point: some auxiliary story must be told to account for the meanings of the primitive expressions, Frege's dictum does not suffice.

A further difficulty is encountered when we reflect on the implications of the principle on which proof-semantics rest. If it is indeed true, as the proof-theoretical semanticist would have it, that speaker S grasps the meaning of the proposition "p" if and only if S knows how to recognize a proof of "p" when one is presented to him, then meaning is *not* individuated per statement, contrary to the molecularist's claim. The ability to recognize a proof of proposition "p" is *not* something specific to "p". When one is in a position to recognize a proof of a statement (of a given language), then it is more than reasonable to expect that he is in a position to recognize proofs for any other statement (of that language). Proof recognition capacities are not specific to a statement but to a language as a whole, so a theory that purports to give meaning to statements *via* these capacities cannot claim that it provides individual meaning, and thus eludes the prospect of holism. It is here that the tension between proof production and proof recognition stands out. On the one hand, in order for meaning to be manifest in use, the focus was put on proof recognition capacities; but on the other, in order to provide an individual meaning to sentences, the focus shifted to canonical proofs, and thus proof production. A canonical proof cannot be the meaning of a statement because it is not always known or even possibly known (cf. [2]), and proof recognition capacities cannot provide individual meaning. Therefore, the antirealist is forced to choose between molecularism and the manifestation thesis.

7.5 The Disputed Class

In his analysis of realism and antirealism, Dummett repeatly claims that "[t]here is no substantial disagreement between the two models of meaning [realism and antirealism] so long as we are dealing with decidable statements" (*TPBOIL*, [5, p. 231]). For a statement D in this class, it would appear then that (1) D has meaning for the realist and antirealist alike and that (2) the realist and antirealist meanings of D are the same. This is a rather peculiar thing to say, because even for statements such as D, the proof of D can be seen as quite a different thing from the truth conditions of D. That is, for the realist at least, a proof of the statement "$1+0=1$" is not the same thing as the fact or state of affairs *one and zero make one*, which is what a he might say the statement "$1 + 0 = 1$" means or corresponds to. To take an example in the common world: a radar gun is the common tool a policeman uses to verify the statement "Jones was driving over the speeding limit", but the realist nonetheless distinguishes the measurement on the radar gun from the state of affairs *Jones was driving over the speeding limit*.

Before getting into the divergence between realism and antirealism on the *un*disputed class, let us try to get the facts straight with the disputed class. Dummett's argument is that truth-conditional semantics fails as a theory of meaning because of the disputed class. Does an antirealist theory of meaning, of the sort envisioned by Dummett, fare any better on the disputed class? We took for granted earlier on the claim that proof-semantics and harmony were enough to guarantee meaning for disputed class statements, but it is now time to examine this in more detail. So what exactly is Dummett's stance on this matter? Let us consider the possibilities I–IV circumscribed in the following table:

ANTIREALISM VS. REALISM ON THE DISPUTED CLASS

	ANTIREALISM	REALISM
Undisputed Statement	*has meaning*	*has meaning*
I. Disputed Statement	*has meaning*	*has meaning*
II. Disputed Statement	*does not have meaning*[a]	*has meaning*
III. Disputed Statement	*has meaning*	*does not have meaning*[b]
IV. Disputed Statement	*does not have meaning*[a]	*does not have meaning*[b]

[a] and he does not claim it does.
[b] but he claim it does.

Clearly, Dummett argues that these two theories differ on the disputed class, so case (I) is automatically eliminated. The reason (IV) is not eliminated, despite the fact that it does not distinguish the realist and the antirealist positions, is because of the slight difference the footnotes *a* and *b* in the table specify: on the antirealist side, it is claimed that a disputed statement has no meaning and the statement, in fact, has no meaning, whereas on the realist side it is claimed that a disputed statement has meaning, but it does not have any (so the realist is wrong in his claim). Since Dummett is implying that a molecular theory of meaning accounts for the meaning of disputed class statements, I take him to be advocating (III) and not (IV), but (II) and (IV) will be used for future reference. Later, we will revisit the claim that a realist theory of meaning fails for disputed class statements (cases (III) and (IV)).

We want to examine if molecularism really holds up to its promises with respect to the disputed class statements; and later, we will examine his further claim that the realist does not have meaning theory for these statements. It can be surprising, at first hand, to think that a theory of meaning based on proof or verification could produce an adequate notion of meaning for undecidable statements, statements which are, by definition, unknowable. But Dummett does not look at the problem from this angle, his focus, rather, is on the notion of conservative extension.

Recall that the molecular Dummett requires that "language as a whole be a conservative extension of that fragment of the language containing observation statements" (*TPBOIL*, [5, p. 221]). This idea has its roots in Hilbert's conception of mathematics. Hilbert thought that only a small fragment of mathematics, "finitisitic" arithmetic, had true content, the rest was just some theoretical apparatus to facilitate investigation of this "contentual" fragment. The idea is expressed by Dummett as follows:

> All other statements of mathematics are devoid of such a content, and serve only as auxiliaries, though psychologically indispensable auxiliaries, to the recognition as correct of the finitistic statements which alone are individually meaningful.[33]

Implicit in this conception is the idea that the *auxiliary* theory must be a conservative extension of the *finitistic* theory, or else the finitistic theory would not provide or contain all truths of contentual arithmetic, and it would therefore be false to call the auxiliary theory an auxiliary to the finististic one. This conception can be compared to an instrumentalist conception of science, where theoretical notions and entities (such as fields and quantas) are introduced only as "fictional" instruments that allow us to predict "real" events in the macroscopic world. So, in the Dummettian picture, disputed class statements are theoretical or fictional statements, and undisputed or decidable statements are the concrete, contentual observation statements.

In the mathematical case, it is somewhat strange to identify the contentual statements with decidable ones. If every instance of an arithmetical predicate $P(x)$ is a contentual statement, that is, if sentences $P(0), P(1), P(2), \ldots$ are all statements of real arithmetic, then I would expect $\forall x P(x)$ to be a statement of "real" arithmetic also, since the attributes "real" and "unreal" qualify the nature of entities the statement is about, not its quantifier complexity. If this is the case, then Gödel's first incompleteness theorem shows us that the undisputed (decidable) class and the class of meaningful, contentual statements do not coincide.

But even if they were to coincide, what kind of meaning would they have on the molecular standpoint? Here is where the conservativeness of the language presumably comes into hand. Disputed class statements contribute to undisputed class statements only as psychological shortcuts: any statement of the undisputed arrived at using disputed class statements could in fact be derived or established with undisputed class statements solely. Hence conservativeness allows the meaning of an undisputed class statement, if its meaning is given in terms of proof, to be given without any disputed class assumptions. What about the meaning of disputed class statements when they are not considered as auxiliaries to real arithmetic? It seems that, on a Dummettian view, we have nothing at all to work with. If knowledge of the meaning of a statement is fleshed out in terms of knowledge of what counts as a (canonical) justification, then unknowable statements, that is to say, statements

---

[33] *TPBOIL*, [5, p. 219].

that lack any canonical justification methods, will turn out meaningless under this account. Our first impressions are confirmed: a verification based theory of meaning fails miserably at providing the class of unverifiable statements with meaning.

There are therefore two "disputed" claims involving molecularism and the disputed class: first, that the undisputed class is well-defined, and second, that conservativeness provides disputed class statements with meaning. Is this restricted to mathematical statements only or can these two claims be disputed in non-mathematical cases too? I would say that things are even worse in the non-mathematical case, where undecidable statements have no logical structure in common.

Consider the following statements, which are all known to be true and involve fairly concrete non-mathematical entities:

(A) Napoleon lost at Waterloo on 18 June 1815.
(B) Sarah Palin was not elected Vice-President of the United States in 2008.
(C) The Andromeda Galaxy can be seen by the naked eye under proper conditions.

I think it is safe to assume that all three statements are decidable. Compare those with the following, which involve the same concrete entities:

(A*) In the evening of 17 June 1815, Napoleon privately thought to himself: "I will lose at Waterloo".
(B*) Sarah Palin will be elected President of the United States in 2012.
(C*) There are humanoid creatures presently living on a certain planet in the Andromeda Galaxy.

If I understand Dummett correctly, they would qualify as undecidable statements (not theoretically but practically), because each concerns either future, past or remote state of affaires of which we cannot know anything. Furthermore, at least one of these statements is undecidable for a rather contingent reason: (A*) could have been decidable had Napoleon left an explicit note to this effect in his diary. Membership to the disputed class is thus sometimes a contingent matter.

The crucial observation is that all three are expressed in the same language as before. Unfortunately, this is bad news for the molecularist, because he will either have to rewrite all statements (A*), (B*) and (C*) in a language $L*$ that is an extension of the language L in which are written (A), (B) and (C), *or* he will have to give up on conservativeness: when two theories are couched in the same language, the only way one can be a conservative extension of the other is if both theories are identical.[34]

There is perhaps a more fundamental reason for these failings of conservativeness, and it comes down to the absence of a systematic account of verification in the

---

[34] For matters of simplicity, "theory" will just mean set of theorems here. Hence, two identical theories can be presented differently.

case of non-mathematical statements. Mathematical statements admit a relatively clean and precise notion of proof or justification: either you have a proof or you don't. I know of no disputes where one mathematician claims a statement to be justified but the other claims it is not and both mathematicians accept the same mathematical principles and rules (and, one might add, level of rigour). There are disputes as to these principles and rules (and to the level of rigour), but that does not keep the intuitionist, for example, from understanding the classicist and vice versa: the intuitionist does not claim that the classical proof is not valid according to classical standards, his claim is rather that classical standards are erroneous. Hence, for a given set of proof standards, justification in mathematics is a fairly circumscribed notion. Furthermore, applying mathematics to the methods of proofs in mathematics, a systematic theory of proof can be given, and it is to this kind of theory that the notion of conservative extension belongs, as well as that of normalisation and canonical proof. As we saw above, with these latter notions, and many more, a precise proof entity (either normalised proof or cut-free sequent proof) can be isolated for a Dummettian-style theory of meaning.

But as the preceding examples tend to illustrate, the picture is not as clear for non-mathematical statements. Can one hope to identify, for each statement, a canonical element of justification? I, for one, am quite sceptical. In an empirical setting, proof is much less tidy. Many reasons explain this. First of all, justification of most statements is of a holistic nature. Even if one does not espouse a fully-fledged Quinian holism, a minimal holism of the sort described by Duhem is hard to refute. Justification for a given physical, chemical, biological or run-of-the-mill down-to-earth statement is multifarious: it can come from different directions, kinds and flavour, and no uniform or canonical notion of justification is available for it. Identifying the logical relations between this statement and others can eventually underline the channels through which credence or confirmation flows from one statement to another, but it will not yield something like a canonical piece of justification.[35]

One must not over-estimate the scope of proof-theoretical considerations of a mathematical kind, and think that they can apply integrally (*mutatis mutandis*) to other areas of discourse. A theory of meaning based on justification will be as good, or bad, as the notion of justification we have at hand. To contrapose: if we have no precise notion of justification (for a given class of statements), we have no such theory of meaning (for that class).

In the end, though, I have trouble understanding why conservativeness would be an enviable property for our language as a whole (with respect to another fragment). If most speakers are like me, they have a fairly limited observation language-theory, and chances are that the language-theory of physics will eventually produce a statement, couched in my observation language, that is not a theorem of my observation theory. To take an age old example, there was certainly nothing in the observation

[35] If one is resolutely opposed to confirmation holism, resorting to induction as a confirmation procedure for general statements will not do either. How many instances of a general law constitute a canonical justification of this law? If a canonical number could be given, induction would never have been a philosophical problem.

language-theory of most astronomers at the beginning of the twentieth century that could predict the bending of light rays by the Sun[36] or, to put it in the observation language of the astronomer, the apparent displacement of stars on the backdrop of the sky close to the Sun during a total eclipse, but that it precisely what general relativity implies. This is certainly not a linguistic or epistemic situation I find objectionable, and if it comes down to choosing between conservativeness and unexpected physical predictions, so much the worse for the former.

7.6 The Obtaining of Truth Conditions

Now that we have seen that Dummett's theory of meaning does not fare as well on disputed class statements, let us see if his claims about the realist meaning theories are themselves justified. It was said that realist truth conditions could not provide disputed class statements with meaning because these conditions would not be manifested in language use. The argument can be represented in condensed form as follows:

1. The meaning of a statement is its truth conditions
2. There are statements whose truth conditions it is impossible to know if they obtain or not
3. Meaning must be learnable, manifest in use
4. $\bot$
5. $\neg$ (The meaning of a statement is its truth conditions)

At this point, it is important to underline that the derivation of the contradiction at step 4 is the result of an important, albeit unmentioned premise, namely:

3.1 If it is impossible to know if the truth conditions of p obtain or not, then it is impossible to know the truth conditions of p

or in contraposed form:

3.2 In order to know the truth conditions of p, one must be able to know if they obtain or not

Only if we add this premise to the assumptions 1, 2 & 3 can a contradiction be derived (and a negation be subsequently introduced on assumption 1). To show you I am being fair with Dummett, and not arbitrarily attributing false premises to him, let me recall one of the (many) passages of *TPBOIL* cited above:

> it is quite obscure in what the knowledge of the condition under which a sentence is true can consist, when that condition is not one which is always capable of being recognised as obtaining.[37]

---

[36] This is actually false. Newtonian physics apparently *does* predict deviations of light rays in this fashion but the deviation predicted is greater than the deviation observed.

[37] *TPBOIL*, [5, p. 225].

The obscurity he is referring to here comes from the fact that if truth conditions must be known (by the manifestation thesis, aka, assumption 3), and if it is impossible to know that certain truth conditions obtain (or not), then what we end up with are unknowable truth conditions that must be knowable, or something like a round square. The careful reader will notice that there is an unmentioned step in this reasoning. The manifestation thesis only requires of truth conditions that they be knowable, *not* that it be possible to know if they obtain or not; if no further assumptions on truth conditions or the learnability of meaning are put forward, the argument produces no contradiction. It is only with the addition of 3.1 (or 3.2) that we obtain the round square. So it appears that in stating his argument Dummett is conflating knowledge of truth conditions and knowledge of the *obtaining* of truth conditions. Both types of knowledge, I would insist, can (and even should) be kept separate.

The advantages of keeping them separate can be observed with the class of counterfactual statements, a class that includes statements reporting dispositions, for the latter are simply disguised counterfactual conditionals. Take, for instance, the statement

(D) Jones is brave

which can be reconstructed as the counterfactual conditional

(D*) Had Jones been in a dangerous situation, he would have acted bravely

where the attribute *bravely* is understood here as non-dispositional. Suppose further that Jones is presently dead, and that he led a quiet and peaceful life, removed from any dangerous situations. In this case, statements (D) or (D*) would be undecidable. The realist, in his view, claims that either "Jones is brave" is true or "Jones is brave" is not true, which is absurd for

> he would then be committed to holding that a statement may be true even though there is nothing whatever such that, if we knew of it, we should count it as evidence or as ground for the truth of the statement.[38]

The reason there is "nothing whatever" is because no situations that could have contributed to our knowledge of Jones' bravery are available to us. (Although we are examining the realist truth-conditional meaning theory at the present, let me briefly come back to the adequacy of the antirealist meaning theory. If statements such as (D) or (D*) are devoid of truth values because there is no way we could put ourselves in a position to know they are true or false, how could they possibly have meaning in a theory where meaning is defined in terms of proof or verification? There simply is no proof or verification that would allow us to assert or deny (D), so would this entail that the statement is meaningless? Furthermore, if a statement lacks meaning, is this not an indication that it is not a statement? Dummett can reject

[38] *Truth*, [5, p. 15].

counterfactual statements if he so desires, but in this case, he cannot claim that his theory of meaning covers disputed class statements.)

I will take for granted that the best truth-conditional theory of meaning presently available for counterfactual statements is the Stalnaker-Lewis account. This account relies on a set of possible worlds W and a system of concentric spheres Σ around each world which is used to express a notion of world-proximity. If "$\varphi > \psi$" represents the conditional "If it were the case that φ, then it would be the case that ψ", then its truth conditions in the Stalnaker-Lewis[39] semantics are

> $w \Vdash \varphi > \psi$ iff either (i) there is no sphere $S \in \Sigma$ containing a φ-world, or (ii) there exists a sphere S such that every φ-world of S is a ψ-world.

The advantage of a systematic account like this one is not so much that it allows us to determine the truth value of counterfactual statements but that it helps us understand, in a systematic way, the possible difference in meaning they may have. Even in the absence of a precise knowledge of the obtaining of counterfactual truth conditions, the formal semantics helps us understands the finer issues that counterfactuals raise.

The fact that possible worlds are presupposed is frequently cited as a weakness of this proposal. Since the non-actual possible worlds are by definition completely disjoint from our reality, in the sense that causal interactions between those worlds and ours (and between the non-actual worlds themselves) are impossible, the critique feels that fundamental tenets of empiricism are being violated and senses the urge to reach for Occam's Browning. But before he unclips his holster, let me remind the critique of the conflation above: there is no reason to suppose that a theory of meaning should provide him with the means to justify assertions.

Does this mean that counterfactuals have meaning but are unverifiable? No, not in the least. It simply means that a theory of justification for such conditionals will not be entailed by the theory of meaning. In my honest opinion, I think theories of justification, for counterfactuals at least, will be harder to come by than theories of meaning. This is mostly for the reasons mentioned previously, namely, that justification is not a well-behaved notion outside mathematical circles. However, I can point to various examples where we do have justification for certain counterfactuals, say:

(E1) If the temperature of the Universe were too high, atoms could not exist;
(E2) Without tides, terrestrial life (on Earth) would not have emerged as quickly;
(E3) Had I not studied in philosophy, I would not be writing this paper.

I consider these three statements to be verifiable and true. How? Not by travelling to the various possible worlds but considering various verifiable truths of the actual world. (E1) is justified by whatever justifies the standard model of particle physics and whatever justifies cosmology (the list of experiments and ancillary hypotheses that one would have to cite are fairly numerous, and I could not name them); (E2) is

---

39 Cf. [11].

justified by a number of general biological principles and by natural selection, whatever justifies them will justify it; and finally, (E3) is justified by various observed tendencies in academic circles plus common sense. All of these elements of justification belong to the actual world, but they nonetheless justify statements whose meanings are spelled out with non-actual worlds.

But what about Jones? What about counterfactuals for which no such justification can be provided? In Jones' case, there is a significant lack of information on the natures of W and Σ (that is, on how Jones would act in the closet possible worlds where he is in a dangerous situation). Here is perhaps where the defender of Dummett's conflation (3.1 & 3.2) could step in, arguing that, if there is no knowledge concerning the obtaining or not of the counterfactual, then there is no way of specifying W and Σ (which allow us to know the truth conditions of the counterfactual). If I *did* know the specific natures of W and Σ, I would know if "Jones is brave" was true or not, provided I knew what world I was in.[40] But then, am I not defending Dummett's conflation? No, not completely. Because even though I do not know what the specific truth conditions are, I know what they look like and what they would need to resemble for (D) to be true or for (D) to be false. My knowledge of the truth conditions allows me to understand what would make a statement like (D) equivalent to or distinct from another statement, despite the relative ignorance.

The situation is similar in game theory. We are unable, in most cases, to specify the exact values of the payoffs of a given game,[41] but general, qualitative information about these payoffs can be sufficient to characterise equilibrium. In cases where our information about "other worlds" is insufficient to determine the truth values of certain counterfactual statements, whatever information we have of other worlds can at least help us understand how these truth values can come to diverge, and more generally, how these truth values happen to be related.

7.7 By Way of Conclusion

I have argued on various levels that a Dummettian meaning theory, or meaning theory template, does not live up to its claims: it does not avoid holism, does not provide disputed class statements with meaning, and does not lead to a rejection of classical logic. Most importantly, it fails to show that truth conditional semantics are unsound. Going back to table 2, the true picture is not so much (III) but a modified version of (II):

---

[40] I do not think that it will do to assume that perfect knowledge of W and Σ is possible *provided that I ignore what world I am in*, I am ready to accept the fact that we do in fact ignore what the specific natures of W and Σ are (if they exist at all).

[41] By this, I mean that the payoffs of *real world* strategic interactions can not always be specified.

ANTIREALISM VS. REALISM ON THE DISPUTED CLASS

	ANTIREALISM	REALISM
Undisputed Statement	*has meaning*	*has meaning*
II.1 Disputed Statement	*does not have meaning*	*has partial meaning*

(The expression "partial" on the realist side is there to underline the fact that, in the case of unverifiable statements, complete knowledge of truth conditions is not possible.)

We should not see truth conditional semantics as a rival to a theory of justification. Proof is our guide to truth, but truth is our guide to understanding what we prove. Both the notion of justification and the notion of truth conditions go hand in hand, one notion is not subordinate to the other.

But then how do I explain that truth conditional semantics commits one to realism and classical logic (whereas justification and proof are neutral on these questions)? Very simply: my explanation is they don't commit us at all to these theses. One can be realist or antirealist about truth conditions. Lewis, for instance, commits himself to the existence of possible worlds, and therefore to the existence of the entities that make up the truth conditions of counterfactual statements. Others take talk of possible worlds, and whatever entity is defined with them, to be just talk. Truth conditional semantics does not come with a specific ontological stance on truth conditions. We can be realists, like Lewis, or antirealists, if, for example, we entertain a formalist or instrumentalist stance on truth conditions. One thing becomes clear however, it is that realism and antirealism are *ontological* not semantic theses, just as they were in the good old days. And as for choice of logic, I would say a similar thing: that truth conditional semantics does not presuppose classical logic, that choice of logic will mostly depend on one's ontological conception of instantiation.

References

1. Belnap, N. 1962. "Tonk, Plonk and Plink." *Analysis* 22:130–34.
2. Boolos, G. 1984. "Don't Eliminate Cut." *Journal of Philosophical Logic* 13:373–78.
3. Boričić, B. R. 1985. "On Sequence-Conclusion Natural Deduction Systems." *Journal of Philosophical Logic* 14:359–77.
4. Burge, T. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4:73–121.
5. Dummett, M. 1978. *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press.
6. Dummett, M. 1991. *The Logical Basis of Metaphysics.* Cambridge, MA: Harvard University Press.
7. Dummett, M. 1993. *The Seas of Language*. Oxford: Clarendon Press.
8. Fodor, J. 1987. *Psychosemantics*. Cambridge, MA: MIT Press.
9. Fodor, J., and Z. Pylyshyn. 1988. "Connectionism and Cognitive Architecture: A Critique." *Cognition* 28:3–71.
10. Gentzen, G. 1969. *The Collected Papers of Gerhard Gentzen,* edited by M. E. Szabo. Amsterdam: North-Holland Publishing Company.
11. Lewis, D. 1973. *Counterfactuals*. Oxford: Blackwell.
12. Prawitz, D. 1965. *Natural Deduction: A Proof-Theoretical Study.* Stockholm: Almqvist and Wilsell.

13. Prawitz, D. 1977. “Meaning and Proofs: On the Conflict Between Classical and Intuitionistic Logic.” *Theoria* XLIII:2–40.
14. Prior, A. N. 1960. “The Runabout Inference-Ticket.” *Analysis* 22:38–9.
15. Read, S. 1999. “Harmony and Automony in Classical Logic.” *Journal of Philosophical Logic* 29:123–54.
16. Tennant, N. 1979. “La barre de Sheffer dans la logique des séquents et des syllogismes.” *Logique et Analyse* 88:503–14.
17. Tennant, N. 1987. *Anti-realism and Logic*. Oxford: Oxford University Press.
18. Tennant, N. 2005. “Rule-Circularity and the Justification of Deduction.” *The Philosophical Quarterly* 55:221.
19. Wittgenstein, L. 1958. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Basil Blackwell.

Chapter 8
Game Semantics and the Manifestation Thesis

Mathieu Marion

8.1 Rethinking the Anti-realist Challenge

Sir Michael Dummett first put forward his ‘anti-realist challenge’ in *Realism*, published in 1963 [27, pp. 145–65]. When reaction came a decade later, it sparked a well-known debate, the ‘realism debate’, that became a chapter in twentieth-century philosophy of language and logic. This debate has now considerably abated, at the advantage of the realist side, where the challenge had not been welcomed to begin with. The initial challenge was for the realist to vindicate her claim that she is in possession of a viable account of meaning for some disputed classes of statements, that would justify adherence to the principle of bivalence – the semantic counterpart to the Law of Excluded Middle—but, on the whole, the reaction had not been to supply such an account, it was rather to reject the challenge itself.[1] It is incumbent on those who feel dissatisfied with the current state of affair, however, to rethink this

[1] Part of negative reaction Dummett’s ideas encountered can be explained by the fact that ‘model-theoretical’ semantics, which takes truth as its fundamental notion and forms the background to realist theories of meaning that he criticized, had been the dominant paradigm in analytic philosophy until then, and has remained so to this day. In my opinion, his ideas were often misunderstood precisely because his critics did not fully understand the ‘proof-theoretic’ background to his ideas. Dummett was rather clear about this on occasions, e.g., when he compared Frege’s axiomatic approach to logic with Gentzen’s proof-theoretic approach and called the former ‘retrograde’ [28, pp. 432f.], or when he modelled the two aspects of language use that he distinguishes on introduction and elimination rules in Gentzen’s natural deduction system; he goes even as far as to generalize the notion of harmony between these, e.g., in [28, pp. 454–55]. It is worth pointing out, however, that arguments such as the Manifestation Argument, discussed here, aim at a particular conception of truth as ‘recognition-transcendent’, not at the concept of truth itself, hence some amount of confusion.

M. Marion (✉)
Canada Research Chair in the Philosophy of Logic and Mathematics, Université du Québec à Montréal, Montréal, QC, Canada
e-mail: marion.mathieu@uqam.ca

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_8,

'anti-realist challenge and to promote some successor challenge or programme.[2] After all, the situation has changed in the intervening half century. At the time Dummett wrote, intuitionistic logic was the only alternative generally considered as a serious rival to classical logic, and he tied his anti-realist challenge at the level of the theory of meaning closely to the adoption of intuitionistic logic, thus renewing the case for the latter at a time when Brouwer's original arguments were simply dismissed as mere 'psychologism' or 'mysticism'. This situation is, however, no longer the case; as the title of a recent collection of essays as it, we now live in *The Age of Alternative Logics* [9]. Other 'deviant'[3] logics, e.g., relevant or paraconsistent systems, have gained enough respect to be counted as genuine alternatives. What this means is that one can no longer feign to ignore them, simply trot out Quine's 'change of language' argument,[4] or limit the debate to classical vs. intuitionistic logic, as Dummett did. Fifty years ago, most philosophers were 'monists', i.e., they believed that there is 'only one logic, today one finds a variety of 'pluralist' programmes.[5] These may either be construed as providing reasons to believe that a plurality of logics are concurrently acceptable, or as providing instead a framework for assessing claims to existence of the various logics now on the market, so to speak.

It is worth exploring how one might want to adjust the anti-realist challenge to this new situation. To put my cards on the table, I am rather inclined to believe, like any good 'monist', that there is only one logic and that it is intuitionistic logic, but I would like a platform from which to assess alternatives. And I think that it is better to explore a possible pluralist framework based on 'game' or 'interaction' semantics,[6] as opposed to better established, more senior 'model-theoretical' or 'proof-theoretical semantics'.[7] In this paper, I want first to suggest how one may get about to rethink the issues raised by Dummett's challenge on this new basis, secondly, to examine closely one of his most important arguments, the Manifestation Argument, and draw some consequences.

The remarks in the next section are meant to give the reader an idea of the form a new programme might take. (One needs in particular to avert misunderstandings,

---

[2] Even if, as I do, one thinks that the adversarial attitude with which the debate had been conducted is mostly out of place. This is also Dummett's opinion [29, p. 464]. Although an adversarial attitude is essential in philosophy, lest we all settle on conceptions that could have been improved upon had they been correctly challenged in the first place, one ought to approach these issues more in the scientific spirit of inquiry, which was that of the pioneers of analytic philosophy, Frege and Russell, than in a quasi-theological spirit.

[3] As W.V. Quine famously called them [66, chap. 6].

[4] For this argument, see [66, p. 81].

[5] To mention some of the better-known proposals: [6], and the programme of 'substructural logics of [24, 71, 79]. Lesser known proposals in [67, 75] are closest in spirit to the standpoint assumed in this paper.

[6] The expression 'semantics of interaction' has the advantage of avoiding some of the inappropriate connotations linked with the concept of 'game'; it was only introduced recently in [48].

[7] So the proposal here is independent from that in [25, 26], further explored in [55, 57] and more closely linked to the 'substructural' programme mentioned in footnote 5.

including a common confusion between 'game' and 'proof-theoretical' semantics.) Detailed explanation are needed but, as space is lacking, my remarks will therefore remain largely suggestive and in need of supplementation at every step: it is hoped that they will have at least introduced a number of key ideas that will come into play in the discussion of the Manifestation Argument.

8.2 Towards a Renewal

The leading ideas of any game semantics are to define logical particles neither in terms of truth conditions,[8] nor in terms of introduction (and elimination rules),[9] but in terms of rules for games between two persons, a proponent and an opponent,[10] and to define logical validity in terms of the existence of a winning strategy for the proponent. Although there are inklings of game semantics in C. S. Peirce's comments on quantifiers, a true game semantics was formulated for the first time by Paul Lorenzen in the late 1950s,[11] and Jaakko Hintikka adapted a variant to fit his model sets a few years later.[12] Lorenzen games were rediscovered and partly adapted by Andreas Blass in 1992 in an endeavour to provide a semantics for linear logic [10]—his paper sparked numerous developments and game semantics quickly reached the status of a 'paradigm' in computer science [78]. These won't be discussed here.[13] Very briefly, and in informal terms, Lorenzen games begin with the proponent P asserting a sentence (or uttering an assertion) and O challenging it, with the players moving in alternate turns. For P as the defender and O as the attacker, the rules that define—or capture—the meaning of the particles are as follows[14]: when P asserts $A \,\&\, B$, O chooses one of the conjuncts and P must defend it, so the game continues for that conjunct; when P asserts $A \vee B$, then O asks that P chooses and defends one of the two disjuncts; for an implication $A \rightarrow B$, O has no choice but concede A in order to force P to defend B; and when P asserts a negation $\neg A$, O has no

---

[8] As is standardly done since, e.g., Wittgenstein's *Tractatus* [90, 5.101]—or in [17, p. 37].

[9] As first done Gentzen's 'Investigations into Logical Deduction', §5.13 [35, p. 80].

[10] More precisely put, these are non-collaborative, zero-sum games with perfect information; there are extensions, e.g., to games of imperfect information, or to games of n-persons.

[11] With further help from his student Kuno Lorenz, in a series of papers reproduced in [51]. For a short introduction, see [74]; for a textbook, in French, see [33].

[12] For one of the first presentations of Hintikka's game-theoretical semantics, which displays the origin in his thinking about model sets, see [41], up to and including Chapter 3. For a more recent detailed presentation, see [46].

[13] In the application of logic to computation and programming language semantics, there are two broad approaches to game semantics, which reflect the approaches to computation as 'proof normalization' and 'proof search'. Game semantics in the style of Lorenzen or Hintikka related more to the later, while developments sparked by [10], such as [1, 3], to be left aside here, relate more to the former. I owe this point to Dale Miller; see [21, §6].

[14] The following are for Lorenzen's games only and, even then, it is necessary at this informal level to gloss over many issues.

choice but to assert A, but then the roles (attack vs. defence) are exchanged, as O has now to defend A against P. For quantifiers, when P asserts $\forall x F(x)$, O chooses a value for x and P must then show that it has the property F, and when P asserts $\exists x F(x)$, then O asks that P exhibits an x that has the property F.

Defined formally, these rules would involve variables for players (X standing for O and P); they would thus be player-independent and come out as symmetric. These particle rules may be said to form a semantic in the sense that they provide an explanation of the meaning of the connectives.[15] In a sense, this semantics is 'local' or 'proto-semantics' [67, p. 366], and needs to be supplemented by some structural rules into a full or 'global' semantics on dialogues such as the obvious rule that players move alternately, one reason being the wish to avoid the 'Tonk' phenomena.[16] It is not my purpose to introduce and discuss structural rules here but some key rules must be mentioned. Here are two: first, it is clear that any such game will end in a finite number of steps, and either P or O will win if and only if it is the other's turn but she cannot move, i.e., either attack or defend. This is the all important Winning Rule.[17] Another key rule, to which I shall come back, is the Formal Rule, which states that P may not introduce atomic formulas, any atomic formulas must be stated by O first. It is crucial since it makes the plays independent of the meaning of the atoms involved in the formulas.[18] With these

[15] One must distinguish here between 'semantics', as understood in this minimal sense and 'model theory': to provide an account of the meaning of logical particles in terms of a key concept (truth, proof, game) is to provide, in one sense of the expression, a 'semantics' for it. (Of course, understood in this minimal sense, semantics need to be supplemented, if only by further 'structural rules'.) And this is quite independent of the fact that one might elect to use model-theoretical tools in fleshing out this semantics, as Hintikka did, for example. For that reason, in his work, game semantics look more like a carrier for the usual model-theoretic notions than a genuine logic of interaction, as pointed out in [2, p. 43, n. 1]. On the other hand, Jean-Yves Girard is quite explicit about avoiding model theory in his highly innovative approach to logic, which he calls 'ludics' [36]. In this paper, there is no need to get into these issues, it suffice for the argument of this paper that one sticks to a minimal definition of semantics as the provision of an account of the meaning of logical particles.

[16] The point is only alluded to in [67, p. 379], for a full discussion, see Rahman's contribution to this volume. For original exchange concerning 'Tonk', see [7, 65].

[17] Another way to frame the Winning Rule is as follows: the game will be over when the last player cannot move anymore; if dialogue is closed, then proponent wins, if it is open, opponent wins. Here 'closed' means that the same formula occurs twice, asserted by O and by P; if not, it is 'open'.

[18] The Formal Rule also introduces asymmetries in the roles of P and O, a fact that has been perceived since [10, p. 185] as a defect. This led to its abandonment within the numerous attempts at giving a game semantics for linear logic, where the above-given particle rules are construed as defining the set of 'additive' connectives, while the new 'multiplicatives' from linear logic are introduced via a new set of rules (the details in this remark are slightly wrong, but this is not the place for a detailed discussion; the gist is, I think, correct). Instead, the essential work done by the Formal Rule is done by the 'copy-cat' strategy. From the Lorenzen standpoint, one might complain that abandoning the Formal Rule brings about some amount of confusion between Lorenzen's games and more model-theoretic versions. Models are indeed often assumed in game semantics for linear logic, e.g., Japaridze's arithmetical models in [47]. And it is a consequence

rules (and others), one can define validity: a formula is valid in a given dialogical system if and only if P has a (formal) winning strategy for this formula, when the Formal Rule is enforced. (Two brief remarks need to be appended here. The first one is needed to avert a possible misunderstanding: although 'global semantics' are kept in the background in what follows, my point is not to focus on particle rules at the expense of structural rules but merely to introduce the minimum necessary for my discussion of the Manifestation Thesis in Section 8.3. Secondly, it is worth noting that tinkering with structural rules while keeping the particle rules unchanged is indeed one way to open the door to pluralism.)[19]

Hintikka's game semantics, which he prefers to calls 'game-theoretical semantics', differs in two fundamental respects. First, there is no rule for $A \rightarrow B$, because Hintikka is a classical logician; for him it is equivalent to $\neg A \vee B$, while in intuitionistic logic the former follows from the latter, but not vice-versa. Secondly, there is no Formal Rule. Instead valuations are introduced as some sort of 'oracle' at the atomic level, and when a play of the game associated to a given formula ends (in a finite number of steps) with one of the players asserting an atomic formula, there is now a different Winning Rule stating that if that player has asserted an atomic formula which is true, then she wins; if the atomic formula is false, then she loses. Of course, in order to find out the truth condition one has to consult the oracle, i.e., look into the model. So Hintikka games are about truth in a model (the main motive was after all to capture Tarski's satisfaction conditions with the existence of a winning strategy), while Lorenzen games are about validity. Lorenzen introduced, however, a rule for 'material games', where the Formal Rule is replaced by a Material Rule stating that atomic formulas standing for true propositions but not false ones may be asserted; the resulting games resemble Hintikka's. It remains, however, that Hintikka's game semantics is model-theoretic and ignores proof theory. The resulting arguments, back and forth, will be by-passed here.[20] The reason for this is that Shahid Rahman and Tero Tulenheimo provided recently a result that paves the way to a better understanding of the relations between the two approaches: they provided an algorithm that transforms winning strategies for Hintikka games into winning strategies for Lorenzen games and vice-versa.[21] I shall briefly come back to this result in the concluding remarks.

These differences between Lorenzen and Hintikka games have important consequences within the 'realism debate', since in the latter there might be facts about the

of this reintroduction of models at the atomic level in absence of the Formal Rule that games are now won or lost for reasons that are, so to speak, independent of the player's moves. Furthermore, asymmetries will resurface in the rules given for the 'multiplicatives'. (I owe these remarks to discussions with Helge Rückert.)

[19] This is the approach taken in [67, 75], which is similar in that respect to the programme of 'substructural logics', mentioned above in footnote 5.

[20] For Hintikka's arguments against Lorenzen, see, e.g., [41, pp. 80–81], [45, p. 39f], [42, pp. 297–98] or [44, p. 267]; for rather poor criticisms of Hintikka from Lorenzen's standpoint, see [32, pp. 352–53].

[21] See Theorem 22 in [68].

model that transcend the players' cognitive abilities and one can prove the existence of a winning strategy which is in principle inaccessible to the players for that reason. This is most obviously the case with 'Henkin' or 'branching' quantifiers, for which the winning strategy will be given by an array of Skolem functions, the existence of which could be proved only classically and of which P would thus have no knowledge; this being an 'existential' fact about the model that obtains independently of the player's abilities to recognize that it obtains or it does not obtain. Hintikka is, of course, quite clear about this[22]:

> The term 'existence' has to be taken seriously here. Strictly speaking, the game-theoretical truth-definition says only that there exists a set of functions, in the most abstract logico-mathematical sense of function, that constitute a winning strategy for the initial verifier. Hence to assert a sentence S is, as far as the basic game-theoretical meaning is concerned, but to make a purely existential statement about the strategies available to the initial verifier in the associated game $G(S)$. Nothing is said whether the player knows what this strategy is or even whether a human initial verifier could know it. For instance, if the winning strategy functions are not computable (recursive), it may be argued that no actual player can play a game in accordance with such a 'strategy'.[23]

This brought about an obvious rejoinder from the anti-realist camp, by Neil Tennant who used a notion of 'strategic intent',[24] in what is but a version of the Manifestation Argument to be discussed below:

> No person could apply these functions in a way that exhibits strategic intent. Nothing the 'possessor' of such a strategy can do could be construed as behaviour manifesting his grasp of such meanings [. . .] We have arrived then, at the following position: in accordance with the Wittgenstenian thesis that one's grasp of meaning must be capable of being manifested eventually and implicitly in observable behaviour, we require the strategies in our game to be effective.[25]

In order to understand better the scope of Tennant's critique, one must distinguish between the *game* itself and *matches* one plays. The analogy with chess might help here: anyone having mastered its rules can go on playing the game of chess, eventually winning matches without possessing an hypothetical winning strategy. The fact that there could be one and that it is unavailable to the players does not bar them from playing, i.e., manifesting their grasp of the rules, independently of their knowledge of a winning strategy. This shows that Tennant's argument applies only at the level of strategies.[26] The result by Rahman and Tulenheimo just mentioned shows, however, that Tennant was not erring in his request that Hintikka games be *effective*, since it shows that they are actually proceeding by an effective construction of the formula.

I have proposed elsewhere another approach to game semantics, in which it is given a philosophical basis in a theory of 'assertion games', which corresponds

---

22 For other relevant passages, see, e.g., [44, pp. 254, 256].

23 [44, p. 171, n. 34].

24 Defined as follows: 'one can behave in such a way as to manifest an intention to effect a certain strategy only if the strategy in question is effective' [87, pp. 304–05].

25 [87, p. 305]. The point is reasserted in [87, p. 176].

26 See [54, pp. 8–9].

roughly to the theory of assertion developed by Robert Brandom's [11, 12, chap. 3], on the basis in particular the chapter on 'Assertion' in Dummett's [28, pp. 295–363].[27] An account of assertion along those lines is certainly one of the many accounts currently available in the market of ideas, so to speak, but this is not to place to defend it.[28] In what follows, I would like (1) briefly to explain the idea, (2) show how it allows one to distinguish game semantics from proof-theoretical semantics, arguing that we should take seriously the dialogical nature of 'assertion games', and (3) push the idea further by associating games with 'proposition'.

1. In *Making It Explicit*, Brandom pictures us as engaged in a perpetual 'game of giving and asking for reasons', hereafter referred to as 'GOGAR'.[29] within which we keep score through 'deontic scoreboards'[30] of each other's 'commitments', i.e. sentences that we committed ourselves to by asserting them or as consequences of sentences one is already committed to, and 'entitlements', i.e., assertions that we have successfully defended in this game [12, chap. 3]. The basis for this approach is a theory of assertions according to which the pragmatic dimension of the speech act of asserting includes the readiness to play this 'game of giving and asking for reasons'. But Brandom, influenced here by Dummett, merely characterized the meaning of logical connectives in proof-theoretical terms,[31] suitably re-described in terms of GOGAR. The suggestion is to look at

[27] See [56] for a brief presentation, and earlier papers [53, 54] where I criticized what I perceived as weaknesses in, respectively, Hintikka's and Lorenzen's attempts at providing a philosophical basis to their own game semantics.

[28] For an overview, see [61]. Pagin has voiced some objections against the 'social account' in [60], but his ultimate claim that assertion would have this particularity that it is the only speech act which is non-social seems hardly credible. This is not the place, however, to discuss it. (John MacFarlane has a rejoinder to Pagin's objections in an unpublished paper entitled 'What is an Assertion?'.) One should note that, as pointed out in [89, pp. 68–69], we are never in a position to defend all of our assertoric commitments. The point about feasibility raised in the papers quoted in footnote 7 above could therefore be adjusted to fit here.

[29] For these games, see also *Between Saying and Doing. Towards an Analytical Pragmatism* [14, p. 111]. But Brandom's project in that book is of a different nature and in order to avoid here needless complications, it shall not be taken into account. The expression 'games of giving and asking for reasons' comes from Brandom, who often attributes it to Wilfrid Sellars, e.g., at [13, p. 189], but, as far as I know, there is only really one passage that supports this attribution, namely the last sentences of §36 of *Empiricism and the Philosophy of Mind* [80, p. 76], where this expression actually does not occur. On the other hand, Dummett came close to framing it when he wrote: 'The process of learning to make assertions, and to understand those of others, involves learning what grounds, short of conclusive grounds, are regarded as justifying the making of an assertion, and learning also the procedure of asking for, and giving, the grounds on which an assertion is made' [28, p. 355]. As it turns out, Dummett used the analogy of chess often, even talking about assertions in terms of 'game', e.g., at [27, p. 2] or [28, pp. 2, 355], but he did not make the notion central to his semantics. Likewise for Wittgenstein, who describes once, almost *en passant*, asserting as 'a move in the language-game' [91, §22]. Therefore, although it is nice to notice affinities, we should avoid reading too much back into the texts of Wittgenstein, Sellars, and Dummett.

[30] See [49].

[31] See his appeal to 'Dummett's Model' in [12, pp. 116–18] and [13, pp. 61–63]. Brandom distinguishes, however, his inferentialism from the sort of 'assertibilism' one finds in Dummett by

GOGAR as the very two-persons games in terms of which logical particles and validity where defined above. An assertion is defined as a move in those games and Lorenzen's dialogical logic is perfectly suited for a precisification of the resulting 'assertion games', since 'asking for reasons' corresponds to 'attacks' in dialogical logic, while 'giving reasons' corresponds to 'defences'. Indeed, 'attacks' are sometimes described as 'rights' while 'defences' are described as 'duties' [50, p. 20], so that we have the following equivalences:

right to attack $\leftrightarrow$ asking for reasons
duty to defend $\leftrightarrow$ giving reasons

The point of winning, i.e., successfully defending one's assertion against an opponent, is that one has thus provided a justification or reason for one's assertion. This gives us the all important philosophical point of playing GOGAR.

2. I would like now to give reasons not to confuse game semantics with the better-known proof-theoretical semantics that forms the basis of Dummett's original anti-realist stance. The central notion is that of 'game', not 'proof' and this is not a superficial change. Keeping the discussion to a strictly philosophical level, one ought to note indeed the fact that, in contrast to dialogical systems, where there are two players, natural deduction systems are 'monological', so to speak, as only one person is involved. Recall that the point of the Formal Rule is that all the elements necessary for P to show the validity of his assertion have to be conceded first by O. The contrast between '*mono*logical' and '*dia*logical' occurs in Brandom's *Making it Explicit* [12, p. 590], but it is construed differently here. Brandom talks of an underlying 'I-Thou' social practice [12, pp. 598–607], but he does not link this to semantics in the same way as proposed here, i.e., through dialogical logic. Instead, he stays close to natural deduction, as we saw [12, pp. 116–18], [13, pp. 61–63]. In his critical study of *Making it Explicit*, Jürgen Habermas has objected on different grounds that Brandom's theory 'does not really do justice to the position of the specific role of the second person' [39, p. 161]. According to Habermas, Brandom never attempts truly to link his theory of assertion with an 'I-Thou' perspective; he keeps instead to a first/third person perspective[32]:

distinguishing two kinds of normative statuses, 'commitment' and 'entitlement', where Dummett would have only one, 'being assertible'. (See [13, p. 188]) The interaction between these gives rise to an incompatibility semantics based on the fact that 'two assertible contents are incompatible in case commitment to one precludes entitlement to the other' [13, p. 194]. This is not the place for further discussion, but one reason one would need to be careful here is simply the fact that details of this incompatibility semantics are worked in a series of appendices to Chapter 5 of *Between Saying and Doing*, the result being that 'any standard incompatibility relation has a logic whose non-modal vocabulary behaves classically' [14, p. 139]. This would prima facie be a problem for someone looking forward to developing a framework for pluralism.

[32] A remark at [12, p. 599] leads me to believe that Brandom was taking his lead from Davidson's 'triangulation'—see, e.g., [23, pp. 117–21, 128–29, 202–03, 212–13]—which requires only two speakers. This is related to but different from what is called here the 'I-Thou' perspective. A set of tangential issues avoided here is raised by Davidson's claims that triangulation is necessary [23, pp. 128–29] and sufficient [23, p. 105] for the *emergence* of thought.

> It is no accident that Brandom prefers to identify the interpreter with a public that assesses the utterance of a speaker—and not with an addressee who is expected to give the speaker an answer. Every round of a new discourse opens with an ascription that the interpreter undertakes from the observer's perspective of a *third* person.[33]

In his reply to Habermas, Brandom conceded the point [13, p. 362], but rejected Habermas' premise that the point of communication is to reach mutual understanding. For him linguistic practice has no particular *raison d'être* [13, p. 363]. This is not the most satisfactory of answers. On the other hand, Habermas' arguments are also not entirely satisfactory, since there are reasons to think that his own position does not truly reflect an 'I-Thou' sociality but privileges instead the perspective of the community, i.e., the 'We'.[34] But he does seem to have a point when he uses Brandom's own example [12, p. 505] of the prosecutor and defence attorney arguing over the trustworthiness of a pathological liar about to take the stand:

> [...] the communicative exchange is played out on two different levels: on one level, both the prosecutor and the defence attorney are speaking to one another in that [...] they reciprocally dispute the correctness of each other's utterance. At the same time, of course, they are aware of the presence of the judges, the jurors, and the public who, on a second level of communication, are following their exchange of words and silently assessing it. Interestingly, Brandom singles out the indirect communication of the speakers with the public who is listening to them—and not the communication of those directly involved—as the paradigm case.[35]

Indeed, one would want here to focus on the prosecutor and defence attorney, and not on their audience; it does not seem right to make the latter play an essential role at the expense of one of the two speakers. The audience's presence is perhaps not entirely insignificant, but it should not play such a central role. The proposal put forward here is to flesh out the theory of assertions in terms of a game semantics instead of 'monological', proof-theoretical terms. It means that the link proposed here is between dialogical logic and 'I-Thou' sociality.[36] To put it crudely, 'I' am P and 'you' are O, so now the Formal Rule states that all elements necessary for I, *qua* P, to show the validity of my assertion have to be conceded first by you, *qua* O. This corresponds fully to the second person standpoint.[37] It is even slightly incorrect to speak as above of the 'interpreter'

---

[33] [39, p. 163].

[34] A reply to Habermas along those lines is found in [77], especially pp. 56–57. Indeed, Brandom considers that there is no 'globally privileged' perspective, including that of the community [12, pp. 599–600], while Habermas' objection stems from the very fact that he considers the community's perspective as privileged: it is for this reason that he misconstrues Brandom as advocating a form of 'methodological individualism' [39, p. 165].

[35] [39, p. 163].

[36] I am aware that this issue between Brandom and Habermas is more complex, since Habermas is discussing here Brandom's peculiar understanding of the *de re/de dicto* distinction, while I keep here to the discussion of logical connectives.

[37] It would be worthwhile to investigate further the connections here with Davidson's 'triangulation' (see footnote 32) and in another direction with Sebastian Rödl's elaboration of the second

as an 'addressee who is expected to give the speaker an answer'; it is better to describe her as a partner in a dialogue whose purpose is to assess the validity of the speaker's claim.[38]

One could point out, however, that this proposed contrast between 'monological' and 'dialogical' is spurious, since one can routinely do proofs in dialogical logic alone, on paper or in one's head. I would be tempted to recall here Sellars' myth of our Rylean ancestors [80, §§48f.], or even Plato's definition of thought as being similar to speech,

> [...] except that what we call thought is speech that occurs without the voice, inside the soul in conversation with itself.[39]

Plato should be understood as stating here not only that thought has the form of a dialogue, but also that it is the 'interiorization' of the external practice of 'dialectical games'.[40] (More on these games below.) The objector would miss here an important philosophical point, which can be put in terms analogous to Dummett's celebrated remark concerning 'judgment' as being the 'interiorization of the external act of assertion' as opposed to the view of 'assertion as the expression of an interior act of judgment' [28, p. 362], by saying that *monological inferential acts are the interiorization of external, dialogical inferential acts*. Note that Aristotle adopted exactly the opposite view in *Posterior Analytics*:

> [...] all syllogism, and therefore a fortiori demonstration, is addressed not to the spoken word, but to the discourse within the soul, and though we can always raise objections to the spoken word, to the inward discourse we cannot always object.[41]

The contrast would therefore be here between a Platonic dialogical model of inferential acts as part of a potentially interiorized dialectical game and a rather Aristotelian monological model of a fully internal series of inferential acts (as in natural deduction) that may or may not be exteriorized.[42] This point might be *philosophical* and might thus make no difference as far as formal matters are concerned, especially given the fact that one can show how strategies for winning dialogues in given systems of dialogical logic correspond to proofs in

person standpoint in [73]. One point especially worth exploring is the essential *symmetry* of the first and second person standpoints, as opposed to the asymmetries of the first and third person standpoints. This essential symmetry is argued for by Collingwood in a splendid passage at [18, pp. 248–49], which was quoted with approval by Davidson [23, p. 219]. This symmetry is, I think, fully cashed out in game semantics by the rule for negation and the possibility to exchange roles.

38 Validity is here understood in the strict logical sense, not in reference to Habermas' wider notion. For more on Habermas' theory see footnote 62 below.

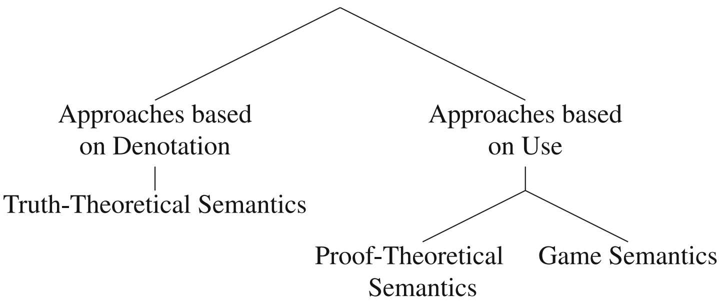
39 *Sophist* 263e; trans. White. For further passages, see *Theaetetus* 189e–190a and *Philebus* 38c–39a.

40 The idea behind this interpretation of Plato's famous saying, I owe to [19, p. 35]. For a similar reading, see [62, pp. 30–33].

41 *Posterior Analytics*, I, 10, 76b 24–27, trans. Mure.

42 Here too, see [62, pp. 36–41]. It is fitting to note here how nicely Aristotle's syllogistic can be expressed within a natural deduction framework, as opposed to an axiomatic one. See [20, 84].

corresponding natural deduction systems,[43] but it is definitely worth making. Some reasons for this will come up later, but one could immediately note here that it affords us a diagnosis of the problem raised by Habermas' critique of Brandom. Although Brandom emphasizes 'I-Thou' sociality, his inferential model is Aristotelian; my suggestion is that the Platonic model is more appropriate. This point is also connected with Hintikka's original motivation for his game semantics, which he took from Wittgenstein's notion of 'language-game': these are rule-governed human activities that serve as 'semantic links between language and reality' [43, pp. 212, 214], i.e., meanings are mediated by language-games.[44] One may fail to be convinced by Hintikka's *Spiel* about games of 'seeking and finding' or by his reading of Wittgenstein,[45] but there seems to be something right with the idea that games are first and foremost an external practice between at least two players. Hintikka often suggested rather implausibly that 'Nature' is to play the role of O, but even if this were the case, it would not make games played between two human players less 'outdoors', contrary to one of Hintikka's early criticisms of Lorenzen [41, pp. 80–81].[46] At all events, if the above is correct, the proper picture of the rival semantic accounts ought therefore to be[47]:



One should now note that the picture was different when the debate surrounding Dummett's anti-realist challenge was raging in the 1970s and 1980s: games semantics was a marginal phenomenon. Lorenzen games and the philosophy of the Erlangen School built around them were simply ignored[48] within 'analytic' circles and Hintikka games, although better known, were hardly mainstream.

---

43 Indeed, since Lorenzen was a 'monist', he believed that his dialogical games would justify intuitionistic logic and this led in his 'Erlangen School' for a search for an equivalence theorem between proofs in Gentzen's natural deduction system for intuitionistic logic and strategies for winning dialogues, and a very 'bureaucratic' proof was given in [31]. See [32] for a readable overview. At all events, a perhaps more convincing argument here would use the fact that game semantics also opens the door to games with n-players, as in, e.g., [2].

44 See, e.g., [43, pp. 151, 156, 174].

45 For criticisms, see [53].

46 See also [67, p. 379].

47 I owe this to Helge Rückert, in a paper given at the Université de Nancy, November 2008.

48 Or summarily dismissed, as in [87, p. 314, n. 6].

There are only a few references to the latter in Dummett's work,[49] and in his own anti-realist camp, only Neil Tennant took them seriously, as we already saw [87, 88]. One should also note that, if the proposed successor programme to Dummett's original 'anti-realist challenge' is to be based on the above theory of 'assertion games', then these programmes will differ in essential ways.

3. Let me push this line of argument a little bit further here by making use of an idea by Aarne Ranta,[50] namely that we see the games that a game semantic will associate to assertions as their 'proposition'[51]: while inferences deal with assertions proper, the content of these is given in terms of the games that the semantic theory associates with them. Here, the rules will capture relevant features of the use of assertions expressing those propositions. This move might not look palatable at first sight, if knowledge of meanings is conceived of as internal and meanings as determinate (more on this last point at the end of next section). It would not be wise, however, to 'intellectualize', in the sense that to 'interiorize' would involve postulating some 'inner process',[52] and conceive of the above particle rules as mental objects, the requirement of which a speaker need to grasp at each new application. In order to do so, one would need to have them available as 'beliefs', but that ought not to be the case. One reason for this can be fleshed out using 'transfer principles', as Stewart Shapiro calls them [81, p. 337]. These are principles such as

$$A, A \to B \vdash B \leftrightarrow (A\&(A \to B)) \to B$$

that establish equivalences between logical inferences and logical truths. But an inference is something someone does, an 'act', while implication is a relation between sentences. Lewis Carroll's paradox in 'What the Tortoise said to Achilles' [15] provided a very good reason not to confuse the two[53]; it is a mistake to think that in order to infer B from the premises A and $A \to B$ one needs to entertain the belief that '$(A\&(A \to B)) \to B$'.[54] One is free therefore to consider instead the idea that inferences such as '$A, A \to B \vdash B$' or

[49] E.g., at [29, p. 84], where Hintikka's game semantics is misconstrued as a species of falsificationism.

[50] See [69, p. 381] and [67, p. 366], where the same line of thought is also pursued.

[51] In the sense of 'proposition' one finds in [17, p. 27], i.e., a proposition is the meaning of a sentence. This is usually meant to be identical with Frege's *Gedanken*, but the notions will diverge sensibly here, as the point is to rethink what 'proposition' might mean. In his paper, Ranta discusses Hintikka games, not Lorenzen's, and his idea paves the way to a suitable reinterpretation of them in type-theoretical terms. But his idea of 'propositions as games' is independent of these issues.

[52] That 'inner process' need not be an 'inner psychological mechanism' of the sort Dummett rejects in [29, p. 37]. The point is fully to adopt the dialogical 'I-Thou' standpoint for which 'know how' is first and foremost of a social practice.

[53] One could also refer here to Wittgenstein's objections in [91, §§185–242], properly understood.

[54] The point is rather standard, see, e.g., [64, p. 291] and [81, p. 337].

'$A \vee B, \neg A \vdash B$' were made before the corresponding beliefs were even explicitly framed. The opposite view would consist in taking seriously the claim that:

> Evidently there must have been a time when the human race, or its immediate ancestor, possessed no logical proposition at all, true or false.[55]

And then claim further that nobody could have actually done a logical inference prior to their corresponding belief having been made explicitly available. But this would be absurd. There would have been, for example, no *elenchus* in Plato's dialogues.[56] The point is, therefore, that the dialogical rules for the connectives are already in use prior to the logician making them explicit when setting up a system which would respect the proprieties of use. So a given speaker may very well have played GOGAR all her life without explicit knowledge of its rules. This point will become important in the next section, when Dummett's use of idea of implicit knowledge to formulate (one of the variants of) the Manifestation Argument will be introduced. The point will be that it is not a speaker's 'implicit' knowledge which is made 'explicit', but the rules of a prior public practice.

To continue references to Plato for the sake of illustration, my claim corresponds to the equation he drew between understanding of the concept 'good', i.e., knowledge of the Form 'the Good', with the ability to win 'dialectical games' about it:

> If someone is incapable of arguing for the separation and distinction of the character of goodness from everything else, and cannot, so to speak, fight all the objections one by one and refute them [. . .], and can't see it all through to the end without his position suffering a fall—if you find someone to be in this state, you'll deny that he has knowledge of goodness itself [. . .].[57]

Plato's dialogues are an excellent source for 'dialectical games', namely a regimented practice of public debate or, to use Habermas' expression, 'ritualized competition for the better arguments' [37, p. 26], which was a common practice in Ancient Greece prior to the introduction of syllogistic by Aristotle in his *Prior Analytics*. These are the games for which Aristotle's *Topics* is the only textbook that survived to this day. These regimented verbal jousts begin with P asserting a claim A, and then, through a chain of questions and answers, in which the adversaries move alternately, O—the role usually played by Socrates—tries to show that the initial claim is part of an inconsistent set of claims $\{A, B_1, \ldots, B_n\}$ held by the proponent—which would form part of what Brandom would call her 'deontic scoreboard'—thus driving her into an elenchus:

$$A, B_1, \ldots, B_n \vdash \perp$$

---

[55] [72, pp. 28–29].

[56] This argument was about inferences but a similar one could be made for particle rules, so it is open to us to conceive of the introduction of rules for the logical particles as 'making explicit' moves already current in a given practice, such as the regimented practice of 'dialectical games' in Ancient Greece and 'obligationes' in the Medieval Ages. (For a game semantic approach to the latter, see [30, part 4].)

[57] *Republic* 534b–d; trans. Grube.

These 'dialectical games' can be shown to proceed according to a set of rules that are closely related to those of Lorenzen's dialogical games, the key difference being that, instead of P arguing for validity, O argues for inconsistency. Therefore, one must introduce for those games a symmetrically related Formal Rule: O may not introduce any thesis A, P must first commit herself to any thesis A; this being related to Socrates' 'disavowal of knowledge', etc.[58] In the passage just quoted, Plato equates 'knowledge of the Good' neither with some propositional knowledge, nor with some intuitive vision of Form but simply with understanding of the concept,[59] this being defined in terms of an ability to remain unbeaten when playing 'dialectical games' over it. (This means being able, when playing O's role, to drive P into an *elenchus*, or, when playing P's role, to avoid being driven into an *elenchus*.) Knowledge of propositional content is thus equated with the ability to play games associated to corresponding assertions. To use a distinction dear to Ryle [76, chap. 2], 'understanding' is here and for Plato very much a matter of 'knowing how', not 'knowing that'. My point is that one can simply see the above set of particle rules as making explicit rules already implicit in practice, i.e., in GOGAR, which forms the context within which the speech act of asserting takes place. Since game semantic is *dia*logical, there is no need to see knowledge of games as some internal 'knowing that'; no 'intellectualization' need be involved.

One last point about Frege's distinction between 'sense' and 'force'. If we are to distinguish between the two, then one should think of the 'theory of force' as providing in the form of rules links between contents and the various speech acts. Limiting myself here to the speech act of asserting, one such rule should look like:

(R) Assert A only if you know a winning strategy for A.[60]

Or, since this rule could only be adhered to by an omniscient being, the weaker rule, perhaps more reasonable in its demands:

(R') Assert A only if you are in a position to defend A.

Presumably, further rules would need to be framed to account for the need to keep one's 'deontic scoreboard' free of contradictions, e.g., the need to 'retract' and 'repair', once one would discover that one's commitment to an assertion A

58 For details about this characterization of 'dialectical games' and the *elenchus* along those lines, see [16]. Note that there is a reference, *en passant*, to the Socratic *elenchus* in [12, p. 178], which shows that the ideas put forth here are not far from Brandom's thinking.

59 For this reading of *Republic* 534b–d, see [5, pp. 283–84]. Of course, the underlying reading of Plato here is very controversial, it certainly has nothing to do with what is peddled as 'Platonism' in the philosophy of logic and mathematics since the early twentieth century but, again, this is not the place for this sort of debate.

60 This is, of course, only a suggestion; the set of such rules would need to be worked out more precisely. Ranta has an initial proposal for such rules in [69, pp. 388 ff.]—in fact, the rule (R) is a variant of one of his rules—but see also his point of departure [85].

is incompatible with, and therefore rules out, entitlement to another assertion B. It is at all events rather important that they form part of the explanation of the very point of playing 'assertion games', in order to complete the semantics and link it with pragmatics; the above Winning Rule being insufficient for this task. As Dummett put it, while discussing the analogy between truth and falsity and winning and losing in a game:

> It is part of the concept of winning a game that a player plays to win, and this part of the concept is not conveyed by a classification of the end positions into winning ones and losing ones. [...] Likewise, it is part of the concept of truth that we aim at making true statements.[61]

This is why Dummett claimed elsewhere that:

> There is a general convention whereby the utterance of a sentence, except in special contexts, is understood as being carried out with the intention of uttering a true sentence.[62]

One may find all manners of defect with this dictum, not least because it involves 'intentions', so one may wish to steer clear of it and aim for rules such as (R) or (R').[63] Such rules square very well with the idea that in order to make a move

---

[61] [27, p. 2].

[62] [28, p. 298]. On the Continental side, Jürgen Habermas developed an 'universal pragmatics' on the basis that a communication is defined as action oriented towards reaching (mutual) understanding. Follows from this what is broadly similar to Dummett's dictum: 'The speaker must have the intention of communicating a true proposition (or propositional content [...]) so that the hearer can share the knowledge of the speaker' [38, p. 22]. In doing so, the speaker raises a 'validity claim' (*Geltungsanspruch*). (For clarification of that notion, see [40].) This means that the speaker is committed to providing reasons for the 'acceptability' of his claim, so: 'We understand a speech act when we know the kinds of reasons that a speaker could provide in order to convince a hearer that he is entitled in the given circumstances to claim validity for his utterance—in short, when we know what makes it acceptable. A speaker, with a validity claim, appeals to a reservoir of potential reasons that he could produce in support of the claim' [38, pp. 232–33]. Habermas' standpoint is thus close, but does not corresponds exactly to Dummett's or Brandom's, it is more like a generalization of it, that embeds it into a larger theory of social action. At all events, it is quite clear, e.g., from [38, pp. 231–32], that it was inspired by Dummett's, and not surprising that Habermas reacted so positively to Brandom's *Making it Explicit*, see [39, chap. 3].

[63] There is also a tangential debate concerning Dummett's mention of 'conventions', since he is generally taken as having fallen foul of a famous argument by Donald Davidson in 'Communication and Convention' [22, pp. 265–80]. (Pagin also picks on Dummett's dictum(s) in [60, pp. 2–3].) Davidson also rejects in his paper the analogy between language and game, which is central here [22, pp. 267–68]. His critique of the analogy consists in the claim that 'linguistic behaviour' does not exhibit a combination of these features of games: (a) people who play want to win, (b) winning is wholly defined by the rules, (c) 'winning can be, and often is an end in itself' [22, p. 267]. He actually concedes partly (b) and provides an obscure Gricean argument concerning (a), that need not be addressed here (as it involves concepts such as 'representing oneself as wanting to win' that are not relevant here). Finally, he claims that 'speaking the truth, in the sense of uttering a true sentence, is never an end in itself' [22, p. 268]. One should note in reply that Davidson was careful enough not to write something like 'winning is always an end in itself', because that would be false; he did not notice that this fact undermined his argument: that speaking the truth is not an end in itself is not contradicting the fact that for some games winning can be an end in itself. The use of

that counts as an assertion, one must have already acknowledged the possibility of a challenge to it.[64]

8.3 The Manifestation Argument and the Manifestation Thesis

In his valedictory lecture delivered in 1992, 'Realism and Anti-realism', Dummett had occasion to revisit the debate generated by his original challenge, and to rectify what he perceived to be misunderstandings, one of which was the idea that he was out to put forward 'a specific thesis of great generality' or the 'platform of a new philosophical party', as opposed to a 'research programme' [29, p. 464]. After all his suggestion had merely been that one could study a variety of traditional disputes, which took the form of a rejection of realism about some classes of putative objects, through a new format, i.e., a debate about a realist theory of meaning that entails, for any such 'disputed class of statements', the adoption of the principle of bivalence, and hence classical logic. One would thus proceed case by case, examining first the reasons why one might adopt a realist theory of meaning for the given class of statements, and possibly argue for a revision of the logic on the basis of anti-realists arguments at the level of the theory of meaning. Of course, the common adversarial response to Dummett was, as he readily admits himself, 'neither wholly right nor wholly wrong'. As he put it:

> What principally interested me, if it did not amount to a single overall thesis, was about a fairly uniform line of argument, not a mere clutch of distinct theses about different subject-matters, united only by bearing a certain structural similarity to one another.[65]

The reasons why critics could be said to be 'neither wholly right nor wholly wrong' won't be discussed here. I merely wish to make a simple, positive suggestion concerning a putative successor programme. A pluralist framework is ideal for a case by case study. When Dummett speaks of a 'fairly uniform line of argument', he probably alludes to the fact, already noted, that any difference between 'realism' and 'anti-realism' about any 'disputed class of statements' would turn around the adoption of the principle of bivalence, so that any two such debates would indeed

the metaphor of games in this paper presupposes that winning is not an end in itself. Indeed, there are many reasons to look at the rule (R) above as involving winning in the assertion game as not being an end in itself. To take only one example, to suppose that O loses to P several plays of, say, a dialectical game, for A, this may convince O that P was correct in asserting A, so that O might not only refrain from further challenges and endorse A. Furthermore, Davidson seems not to have realized that Dummett's point remains even when the analogy does not hold: one needs an account of what it is that a speaker aims at when 'asserting A' for one's semantic account to be complete.

[64] Shieh's reformulation of Dummett's dictum, as follows, comes rather close to the position advocated here: 'To be taken as making an assertion, a speaker must acknowledge that the statement she is making is subject to assessment as correct or incorrect, by reference to what she would count as justifying it' [82, p. 51].

[65] [29, p. 464].

have a fair deal in common. There is a risk that any such 'fairly uniform line of argument' might be jettisoned in a pluralist framework, where room would be made for other possible revisions of classical logic than the specific one advocated by intuitionists. This, however, would cause some problems for anyone wanting to put forth a successor programme to Dummett's anti-realist challenge. In other words, what would happen of the arguments put forth by anti-realists in support of their (intuitionistic) revision of classical logic?

The Manifestation Argument is one such argument. It is found in many places, e.g., in the opening pages of 'The Philosophical Basis of Intuitionistic Logic' [27, pp. 215–47], in Chapter 13 of *Frege. Philosophy of Language* [28, pp. 466–68] or in 'What is a Theory of Meaning? (II)' [29, pp. 46–47], where it is applied to various classes of statements, from mathematics to conditionals. This argument is notoriously difficult to pin down, not only because one can extract from Dummett's text various distinct arguments, but also because influential critics, such as Colin McGinn and John McDowell, or even allies, such as Neil Tennant, misrepresented it to some extent.[66] In this section, I shall try and clarify the nature of the argument, making use of a pair of excellent papers by Sanford Shieh [82, 83].

It will be useful to get some definitions across first. The theory of meaning that Dummett argues against has two features. First, its key semantic concept is 'truth':

(1) The meaning of a declarative, non-indexical sentence is the condition under which it is true.

Secondly, the concept in question is that of 'recognition-transcendent' truth, which could be framed as follows:

(2) The truth-condition of an undecidable sentence can obtain or fail to obtain independently of our capacity, even in principle, to recognize that it obtains or fails to obtain.

It goes without saying that 'recognition-transcendence' is postulated here to make sure that concept of truth in (1) satisfies the principle of bivalence (every meaningful declarative sentence is either true or false). Thesis (2) relies on a distinction between 'decidable' and 'undecidable' sentences, which is peculiar to Dummett. These notions should not to be confused with their ordinary counterparts in mathematical logic; there is an obvious reason for this, as Dummett needs a distinction that would be serviceable in other contexts than mathematics, a distinction closely linked to the idea of 'recognition-transcendence'. Dummett is quite clear about this:

> Many features of natural language contribute to the formation of sentences not in principle decidable: the use of quantification over an infinite or unsurveyable domain (e.g., over all

[66] See [58, 59], and, on the anti-realist side [88]. The first two are soundly criticized in [82] for the quasi-behaviouristic construal of the manifestation argument (see footnote 85 below), which is neither really compelling, nor faithful to Dummett. (The latter's formulations are admittedly often ambiguous, given the context in which they were written, permeated as it was by Quine's philosophy.) But it turns Dummett into some sort of easily refutable old-style 'epistemological foundationalist'.

> future times); the use of the subjunctive conditional, or of expressions explainable only by means of it; the possibility of referring to regions of space-time in principle inaccessible to us.[67]

I reproduce here Shieh's definitions:

(3) A sentence is *decidable* just in case we have already provided a proof or a refutation of it or we are in possession of an effective method for it. In other words, a sentence is decidable just in case we know or have reason to think that either we can recognize it to be true or we can recognize it to be false.

(4) A sentence is *undecidable* just in case it is not decidable, that is that neither do we know nor do we have reason to think that either we can recognize it to be true or we can recognize it to be false [83].

That such a notion had to be introduced is a requirement of Dummett's attempt at forging a 'fairly uniform line of argument', as explained above. It is not clear, however, if this notion of 'undecidability' is ultimately coherent or not.[68] At all events, the point of those definitions is that Dummett understands his quarrel with the realist to be concerned only with 'undecidable' sentences. He concedes that there is no grounds for disagreement with the realist about the meaning 'decidable' sentences, so his argument is tailored to be effective only in the case of 'undecidable' sentences:

> When the sentence is one which we have a method for effectively deciding, there is again no problem: a grasp of the condition under which the sentence is true may be said to be manifested by a mastery of the decision procedure, for the individual may, by that means, get himself into a position in which he can recognize that the conditions for the truth of the sentence obtains or does not obtain, and we may reasonably suppose that, in this position, he displays by his linguistic behaviour his recognition that the sentence is, respectively, true or false. But, when the sentence is one which is not [. . .] effectively decidable, as is the case with the vast majority of sentences of any interesting mathematical theory, [. . .] the condition which must, in general, obtain for it to be true is not one which we are capable of recognizing whenever it obtains, or of getting ourselves in a position to do so.[69]

Finally, one needs a Manifestation Thesis; here is one of Dummett's versions of it:

(5) Manifestation Thesis: 'if two individuals agree completely about the use to be made of the statement, then they agree about its meaning'.[70]

A version for which he gives the following motivation in terms of chess, i.e., in implicitly game semantical terms:

> The reason is that the meaning of a statement consist solely in its role as an instrument of communication between individuals, just as the powers of a chess-piece consist solely in its role in the game according to the rules.[71]

67 [29, p. 46].

68 Despite his best efforts, Shieh concludes [83] on a sceptical note.

69 [27, p. 225].

70 [27, p. 216], [63, p. 4].

71 [27, p. 216].

Given these ingredients, the Manifestation Argument can now be stated. There are three variants that can be extracted, about conveyability, implicit knowledge and acquisition or learning; I shall not discuss the last one.[72] The Conveyability Argument, as found in 'The Philosophical Basis of Intuitionistic Logic' goes as follows:

(6) Given the Manifestation Thesis, meaning is exhausted in use. Meaning cannot 'contain as ingredient, anything which is not manifest in the use made of it, lying solely in the mind of the individual who apprehends that meaning'.[73]
(7) Therefore, a speaker cannot be said to communicate more than what he is observed to communicate.[74]
(8) Now, a speaker's capacity for acknowledging sentences as true (or false) can be displayed only in cases in which the conditions for their truth can be recognized as obtaining (or failing to obtain), i.e., for 'decidable' sentences.
(9) Given that in the case of 'undecidable' sentences a speaker cannot, even in principle, recognize that the conditions for their truth obtain (or fail to obtain), what the speaker is observed to communicate cannot be a full manifestation of the speaker's knowledge of these truth-conditions.[75]
(10) Therefore, a theory whose features are (1)–(2) above cannot provide an account of 'undecidable' sentences; it 'cannot be a theory in which meaning is fully determined by use' [27, p. 225].

The Implicit Knowledge Argument requires further ingredients, starting with Dummett's controversial thesis that:

(11) Understanding is knowledge of meaning.[76]

Because of this, Dummett cannot be satisfied with the above distinction between 'sense' and 'force', and needs to supplement the theory of meaning with a third part, which would specify 'in what having this knowledge consists, i.e., what counts as a manifestation of that knowledge'.[77] For model-theoretic semantics based on the notion of truth condition, Dummett prefers to speak of 'theory of reference' instead of 'theory of sense' [29, p. 40], so we can use this last expression for his idea of a

---

[72] See [27, pp. 217–18].
[73] [27, p. 216].
[74] [27, p. 216].
[75] The following passages support this reconstruction: 'it is quite obscure in what the knowledge of the conditions under which a sentence is true can consist, when that condition is not one which is always capable of being recognized as obtaining' [27, p. 224]; 'any behaviour which displays a capacity for acknowledging the sentence as being true in all cases in which the condition for its truth can be recognized as obtaining will fall short of being a full manifestation of the knowledge of the conditions for its truth: it shows only that the condition can be recognized in certain cases, not that we have a grasp of what, in general, it is for that condition to obtain even in those cases when we are incapable of recognizing that it does' [27, p. 225].
[76] [29, pp. 3, 35].
[77] [29, p. 37]. For this tripartite conception of the 'theory of meaning', see [86, §3].

third part. Two further remarks needed to be added here about this 'theory of sense'. First, there is the idea of 'tacit' or 'implicit' knowledge:

(12) Knowledge of meaning can only be implicit.[78]

Dummett motivates the thesis as follows:

> In general, it cannot be demanded of someone who has any given practical ability that he have more than an implicit knowledge of those propositions by means of which we give a theoretical representation of that ability. [...] It would, moreover, be palpably incorrect to hold that, once someone had mastered a language, he could give, in that language or any other, an explicit formulation of a theory of meaning for that language.[79]

It is worth interjecting here a critical comment. Remarks such as these imply that Dummett conceived the theory of meaning as providing a 'theoretical representation', which could not be realistically said to be possessed by anyone having mastered a language, so that knowledge of meaning could only be 'implicit'. Remarks in the same passage lead one to believe that Dummett had in mind, strangely enough, an *axiomatic* theory of meaning [29, pp. 37–38],[80] in which case it is quite obvious that knowledge of it cannot be explicit. A subtle shift seems to have occurred here since the point of a theory of meaning was to capture abilities manifested by a speaker in a linguistic practice she finds herself already engaged into, not to try and provide a 'theoretical representation' of some mental capacity supposedly needed to explain a speaker's linguistic behaviour. That the theory is to saddle the speaker with a 'theoretical representation' that she cannot handle should not be seen as an indication that this 'theoretical representation' of her abilities must be implicit, but that the idea might rather be misguided in the first place: it is not a speaker's interiorized *implicit* knowledge that the theory is supposed to *make explicit*, but the (set of) rules of a prior public practice in which she is engaged. In other words, the theory should aim at capturing the set of rules that determine moves available to speakers, not beliefs (e.g., axioms and theorems) the possession of which by speakers being implausibly postulated by the theory.

Secondly, if the theory is a realist one, i.e., with features (1)–(2) above, then this 'theory of sense' will have to state how one fully manifests knowledge of truth-conditions that might possibly be recognition-transcendent. With this we can state the Implicit Knowledge Argument, as found in 'What is a Theory of Knowledge? (II)'[81]:

(13) In the case of 'decidable' sentences, a speaker's knowledge of the conditions for them to be true consists in a practical ability, namely 'his mastery of the procedure for deciding it, that is, his ability, under suitable prompting, and

[78] [29, p. 36]. To advert a misunderstanding, the notion of 'implicit knowledge' discussed here is not that of [8]. For a discussion of game semantics in terms of the latter, see [70].

[79] [29, pp. 36–37].

[80] Compare [27, p. 451]. See also footnote 1 above.

[81] See also [27, p. 217].

display, at the end of it, his recognition that the condition does, or does not, obtain'.[82]

(14) In the case of 'undecidable' sentences, i.e., sentences with recognition-transcendent truth-conditions, 'we cannot equate', by definition, 'a capacity to recognize the satisfaction or non-satisfaction of the condition for the sentence to be true with a knowledge of what that condition is'.[83]

(15) Therefore, a speaker's knowledge of meaning, understood as knowledge of the condition a sentence to be true, cannot be fully manifested in a practical ability. The theory of sense is incomplete.

With hindsight, an obvious comment on the Conveyability and Implicit Knowledge Arguments is that they rely on a number of contentious notions, e.g., that of an 'undecidable' sentence or that of 'implicit knowledge', with the concomitant need for a theory of meaning to be either 'modest' or 'full blooded', notions about which quite a lot of ink has been spilled. One would be ill-advised to get embroiled in these old debates.[84] Perhaps one could simply separate the Manifestation Thesis, which is much more intuitively acceptable, from the Manifestation Argument in which it is embedded – as can be seen in the two variants just discussed. This move would have the advantage of separating the wheat from the chaff, so to speak, since one does not need anymore to appeal to these controversial notions.

This would still leave us, however, with the fact that, as mentioned earlier, both arguments are likely to be misinterpreted as reflecting misplaced worries about the epistemology of communication: it looks as if according to Dummett's conception meanings are somehow hidden from view, as if we would have no access to them if they were not manifested in some 'observable behaviour'.[85] Sanford Shieh has developed an alternative reading—to my mind a much more interesting one—according to which the claim is instead that:

> We have *no coherent conception of there being such a thing as the meaning someone associates with a statement*, unless she manifests it in what she counts as justifying the statement. The philosophical basis of the [Conveyability Argument] can now be seen as antecedent to the epistemology of meaning.[86]

---

[82] [29, pp. 45–46].

[83] [29, p. 46].

[84] This might explain why Ranta chose simply to by-pass the Manifestation Argument in [69, p. 386].

[85] The standard 'behaviourist' readings of the Manifestation Argument consists in transcribing it in terms of Quine's 'dispositions to verbal behaviour', e.g., dispositions to assent or dissent from a sentence when its truth condition obtains or fails to obtain. See [82, pp. 38f.]. There are passages where Dummett criticizes Quine's impoverished conception of language use, e.g., [28, p. 614], that should alert one to the fact that this transcription could not be faithful to Dummett's views. Furthermore, dispositions to assent or dissent are meant to provide a valuation of atomic propositions, they could not be what Dummett, who was consciously shedding valuations, was looking for.

[86] [82, p. 61].

A way out of these difficulties is afforded by a closer look at the Manifestation Thesis (5) and a possible argument in its favour. In accordance with his interpretation, Shieh suggested a variant:

(5′) Manifestation Thesis: 'if two speakers agree in what they would *count* as justifications for a statement, then they attach the same meaning to it'.[87]

I would suggest that one takes seriously the fact that in (5) and (5′) one is talking about *two speakers*. Although hardly ever noticed, this fact is far from incidental. The 'I-Thou' perspective is indeed, upon consideration, essential to the very formulation of the Manifestation Thesis, and the proposal under consideration in this paper is to capture it through a game semantics such as the above, where agreement on justifications can be seen as obtained through playing GOGAR. Furthermore, this fits nicely the idea, also introduced above, of 'games as propositions', if we are to see the games associated with propositions as a two-persons social practice, i.e., as dialogical games, as opposed to interiorized monological games. Assuming that the speaker whose claim has been put forward is to play the role of P and the hearer that of O, working towards an agreement on what would *count* as justifications for a given claim is a collaborative endeavour involving both P and O. The latter is not a passive third-person audience, she is rather involved in an essential way, as we saw from consideration of the meaning of the Formal Rule. The viewpoint presented in this paper would thus be that the Manifestation Thesis (5′) is implicitly presupposed by the very idea of game semantics.

Therefore, in the kind of successor programme envisaged in the previous section, based on game semantics, the Manifestation Thesis would not play the same role. It would merely have now an explicatory role within the philosophical basis of the proposed game semantics: two speakers attach the same meaning to an expression if they agree in what they would *count* as justifications for it; this agreement being reached precisely by playing GOGAR. The reason for this is that meaning is tied in GOGAR to *the ability manifested in playing the game*, the same way that in chess understanding of the rules of the game—and therefore knowledge of the meaning of the pieces—is displayed in playing chess, with the difference here that it would be like picking up how to play chess without being told the rules, as these would be made explicit at a later stage. What is displayed by any competent chess player is completely displayed in her play: there is no ingredient missing, as if one would also need knowledge of a winning strategy in order really to be said to know chess.

To illustrate again with the above 'dialectical games', that P drives O into an *elenchus* shows that O needs to repair his 'deontic scoreboard'; it also shows a discrepancy between the players' understanding of the concept or statement involved, at the advantage of P, since in these games it is O who is shown to have a potentially defective understanding. For example, in Book I of *Republic*, Socrates, as O, drives twice Thrasymachus, as P, into *elenchi*, at *Republic* 343a and 354a, thus

[87] [82, p. 50]. One may wish to speak here more generally about 'expressions', instead of 'statements'.

showing with these two matches that Thrasymachus' understanding of 'justice' is defective.[88]

As the example of 'dialectical games' in Plato's dialogues show, however, it is not clear that one could play such games or GOGAR to a satisfactory degree, i.e., to a point where complete agreement is reached. This raises a related issue concerning the underlying notion of meaning. It is not clear, at least at first blush, that on the conception advocated here one is committed to a notion of 'determinate meaning', in the tradition of Frege and the early Wittgenstein; one ought perhaps talk instead of a 'theory of meaningfulness',[89] according to which dialogue partners seek understanding only to the extent that it allows for coordination aiming at some common task, as opposed to a 'theory of meaning' that attempts to provide a theoretical representation of an implicit knowledge of 'determinate meanings'.

8.4 Concluding Remarks

If one separates the Manifestation Thesis from the Manifestation Argument, what remains of use of the latter as a 'fairly uniform line of argument', within the 'anti-realist challenge'? I would like to conclude with some further speculative remarks on this point. Recall that the overall plan was to try and provide a successor to the programme involved in Dummett's original challenge, one for which pluralism is a desideratum. Furthermore, the (provisional) idea was that the plurality of logics would be captured by keeping the particle rules constant while varying the structural rules on games; the debates would then be on the acceptability of the structural rules needed for this or that logic. If the above is allowed to stand, then the anti-realist is not going to appreciate the following consequence: if the Manifestation Thesis is to play within this context the role I suggest it should, then it is independent of any choice of logic and a classical logician would not have to worry about it, her understanding of the connectives would equally be manifest in her use of them in games with (structural) rules for classical logic. So *the Manifestation Thesis would cut no ice* and Dummett's 'anti-realist challenge' would be severely undermined.

However, a reluctant realist wishing to adhere to a 'recognition-transcendent' concept of truth might object to game semantics to begin with, arguing that it is merely 'verificationist' and that 'verificationism' cannot issue in any recognizable form of realism. It is true that Dummett presented his whole approach in terms of

---

88 The example of 'dialectical games' is particularly apt, since they were played before Aristotle ushered logic by introducing his syllogistic in *Prior Analytics*. The point is historical but serves to illustrate that the introduction of logical rules can be seen as *making explicit* rules already in use in a practice. What made Athens in the fourth century BCE peculiar was the regimentation of those 'dialectical games', which forced one to provide training for them in the form of recipes, that were eventually to become Aristotle's syllogistic.

89 I got this notion from reading [77].

'verificationism'.[90] Witness this passage from 'What is a Theory of Knowledge? (II)':

> I have argued that a theory of meaning in terms if truth-conditions cannot give an intelligible account of a speaker's mastery of his language; and I have sketched one possible alternative, a generalization of the intuitionistic theory of meaning for the language of mathematics, which takes verification and falsification as its central notions in place of those of truth and falsity.[91]

It is also true that Hintikka argued that his game semantics is in effect a 'synthesis' of verificationist and truth-conditional semantics.[92] After all, as Hintikka pointed out, in his own game-theoretical semantics

> [. . .] truth-conditions are formulated by references to certain games. These truth-conditions are therefore understood by a speaker as soon as he or she understands these games. You can give this understanding as radically Wittgensteinian a turn as Dummett might wish, but you cannot escape admitting that a speaker's knowledge of truth-conditions is manifested as completely as any tacit knowledge can ever be manifested in his or her mastery of semantic games. Such a mastery can be complete without its resulting in a decision method for truth and falsity, or even in a definitive decision concerning truth or falsity in the case of particular given theories.[93]

Taken as applying generally to game semantics, his point is exactly that of this paper. This can nevertheless still be taken as an indication that game semantics is already biased towards anti-realism; and our reluctant realist could still argue that, if it is true that game semantics is verificationist, then game semantics can't be truly realist. Hintikka would beg to disagree, pointing out that:

> There is a tacit assumption in Dummett and in several other recent philosophers which is refuted by game-theoretical semantics. It is that any reliance on human activities which serve to bridge the gap between language and the world is inevitably a step from realism.[94]

The result by Rahman and Tulenheimo mentioned above might help us see our way through this. It shows that Hintikka and Lorenzen games proceed likewise, by an effective construction of the formula. So Hintikka games are in fact intuitionistic in the way they proceed,[95] but they remain classical only because his models are such

---

[90] It should be understood, of course, that what is at stake here is not a revival of the 'verificationism' of the logical positivists.

[91] [29, p. 74].

[92] See [44, pp. 250–73] for a paper with this very title.

[93] [44, p. 255]. The point expressed here in Hintikka's last sentence is quite correct, inasmuch that the contrary claim would embroil its supporters in ridiculous claims about the fact that chess players would not know how to play chess unless they would know a winning strategy for all possible configurations of the board. Winning strategies are known for some simple configurations close to check mate, but it would be ridiculous to claim that one who does not know them does not know how to play chess.

[94] [44, p. 257].

[95] For example, when in a Hintikka game P has to choose which of $A \vee B$ she should defend, she gets her answer from the model. This corresponds to the intuitionist notion of disjunction.

that atomic formulas satisfy the Law of Excluded Middle.[96] It may therefore be that it is the realist who won't accept the move to game semantics, but it should not be too difficult to give reasons why models should satisfy the Law of Excluded Middle. To carry the discussion further would require another paper, but at least this is clear for now: the 'realism debate' is not closed with the move to game semantics, even if the sting has been taken out of the Manifestation Thesis. It is merely taking another form. Let me simply emphasize in closing the tentative nature of these last remarks. It was not my brief to argue for one side of the 'realism debate' here, but simply to outline a successor programme and argue that the Manifestation Thesis naturally fits it.

Acknowledgements Many thanks to Aude Bandini, Benoît Castelnérac and, especially, Helge Rückert for helpful conversations on points raised in the paper, and Shahid Rahman for comments on an earlier version.

References

1. Abramsky, S. 1997. "Semantics of Interaction: An Introduction to Game Semantics." In *Semantics and Logics of Computation*, edited by A. M. Pitts and P. Dybjer, 1–31. Cambridge, MA: Cambridge University Press.
2. Abramsky, S. 2006. "Socially Responsive, Environmentally Friendly Logic." In *Truth and Games. Essays in Honour of Gabriel Sandu*, edited by T. Aho and A.-V. Pietarinen, 17–45. Helsinki: Societas Philosophicas Fennica.
3. Abramsky, S., and R. Jagadeesan. 1994. "Games and Full Completeness for Multiplicative Linear Logic." *Journal of Symbolic Logic* 59:543–74.
4. Aho, T., and A.-V. Pietarinen, eds. 2006. *Truth and Games. Essays in Honour of Gabriel Sandu*. Special issue of *Acta Philosophica Fennica* 78.
5. Annas, J. 1981. *An Introduction to Plato's Republic*. Oxford: Clarendon Press.
6. Beall, J. C., and G. Restall. 2006. *Logical Pluralism*. Oxford: Clarendon Press.
7. Belnap, N. 1962. "Tonk, Plonk and Plink." *Analysis* 22:130–34.
8. van Benthem, J. 1991. "Reflections on Epistemic Logic." *Logique et Analyse* 133–134:5–14.
9. van Benthem, J., G. Heinzmann, M. Rebuschi, and H. Visser, eds. 2006. *The Age of Alternative Logics. Assessing Philosophy of Logic and Mathematics Today*. Dordrecht: Springer.
10. Blass, A. 1992. "A Game Semantics for Linear Logic." *Annals of Pure and Applied Logic* 56:183–220.
11. Brandom, R. 1983. "Asserting." *Noûs* 17:637–40.
12. Brandom, R. 1994. *Making It Explicit. Reasoning, Representing & Discursive Commitment*. Cambridge, MA: Harvard University Press.
13. Brandom, R. 2000. *Articulating Reasons. An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.
14. Brandom, R. 2008. *Between Saying and Doing. Towards an Analytic Pragmatism*. Oxford: Oxford University Press.
15. Carroll, L. 1895. "What the Tortoise Said to Achilles." *Mind n.s.*, 4: 278–80.
16. Castelnérac, B., and M. Marion. 2009. "Arguing for Inconsistency: Dialectical Games in the Academy." In *Acts of Knowledge: History, Philosophy and Logic. Essays Dedicated to Göran Sundholm*, edited by G. Primiero and S. Rahman, 37–76. London: College Publication.

96 See [68, p. 203] and [48, pp. 158–59].

17. Church, A. 1956. *Introduction to Mathematical Logic*. Princeton, NJ: Princeton University Press.
18. Collingwood, R. G. 1938. *The Principles of Art*. Oxford: Clarendon Press.
19. Collingwood, R. G. 1939. *An Autobiography*. Oxford: Clarendon Press.
20. Corcoran, J. 1974. "Aristotle's Natural Deductive System." In *Ancient Logic and Its Modern Interpretations*, edited by J. Corcoran, 85–131. Dordrecht: D. Reidel.
21. di Cosmo, R., and D. Miller. 2010. "Linear Logic." In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. http://plato.stanford.edu/entries/logic-linear/#ComSciImp/.
22. Davidson, D. 1984. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
23. Davidson, D. 2001. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
24. Dosen, K. 1989 "Logical Constants as Punctuation Marks." *Notre Dame Journal of Formal Logic* 30:362–81.
25. Dubucs, J. 2002. "Feasibility in Logic." *Synthese* 132:213–37.
26. Dubucs, J., and M. Marion. 2003. "Radical Anti-realism and Substructural Logics." In *Philosophical Dimensions of Science. Selected Contributed Papers from the 11th International Congress of Logic. Methodology, and the Philosophy of Science, Krakow, 1999*, edited by A. Rojszczak†, J. Cachro, and G. Kurczewski, 235–49. Dodrecht: Kluwer.
27. Dummett, M. A. E. 1978. *Truth and Other Enigmas*. London: Duckworth.
28. Dummett, M. A. E. 1981. *Frege. Philosophy of Language*. London: Duckworth.
29. Dummett, M. A. E. 1993. *The Seas of Language*. Oxford: Clarendon Press.
30. Dutilh Novaes, C. 2007. *Formalizing Medieval Logical Theories. Suppositio, Consequentiae and Obligationes*. Dordrecht: Springer.
31. Felscher, W. 1985. "Dialogues, Strategies, and Intuitionistic Provability." *Annals of Pure and Applied Logic* 28:217–54.
32. Felscher, W. 1986. "Dialogues as a Foundation for Intuitionistic Logic." In *Handbook of Philosophical Logic*, edited by D. Gabbay and F. Guenthner, vol. III, 341–72. Dordrecht: D. Reidel.
33. Fontaine, M., and J. Redmond. 2008. *Logique Dialogique: une introduction, Volume 1: Méthode de Dialogique: Régles et Exercices*. London: College Publications.
34. Gabbay, D., and F. Guenthner, eds. 1986. *Handbook of Philosophical Logic*, vol. III. Dordrecht: D. Reidel.
35. Gentzen, G. 1969. *The Collected Papers of Gerhard Gentzen*, edited by M. E. Szabo. Amsterdam: North Holland.
36. Girard, J.-L. 2001. "Locus Solum." *Mathematical Structures in Computer Science* 11:301–506.
37. Habermas, J. 1981. *The Theory of Communicative Action*, vol. 1. Boston, MA: Beacon Press.
38. Habermas, J. 1998. *On the Pragmatics of Communication*. Cambridge, MA: MIT Press.
39. Habermas, J. 2005. *Truth and Justification*. Cambridge, MA: MIT Press.
40. Heath, J. 1998. "What Is a Validity Claim?" *Philosophy and Social Cricitism* 24/4:23–41.
41. Hintikka, J. 1973. *Logic, Language Games, and Information*. Oxford: Oxford University Press.
42. Hintikka, J. 1987. "Replies and Comments." In *Jaakko Hintikka*, edited by R. Bogdan, 277–344. Dordrecht: D. Reidel.
43. Hintikka, J. 1996. *Selected Papers Volume 1, Ludwig Wittgenstein: Half Truths and One-and-a- Half Truth*. Dordrecht: Kluwer.
44. Hintikka, J. 1998. *Selected Papers Volume 4. Paradigms for Language Theory and other Essays*. Dordrecht: Kluwer.
45. Hintikka, J., and J. Kulas. 1985. *The Game of Language. Studies in Game-Theoretical Semantics and its Applications*. Dordrecht: D. Reidel.
46. Hintikka, J., and G. Sandu. 1997. "Game-Theoretical Semantics." In *Handbook of Logic and Language*, edited by J. van Benthem and A. ter Meulen, 361–410. Amsterdam: Elsevier.
47. Japaridze, G. 1997. "A Constructive Game Semantics for the Language of Linear Logic." *Annals of Pure and Applied Logic* 85:87–156.

48. Keiff, L., and S. Rahman. 2010. "La dialectique, entre logique et rhétorique." *Revue de métaphysique et de morale* (2):149–78.
49. Lewis, D. 1983. "Scorekeeping in a Language Game." In *Philosophical Papers*, vol. I, 233–49. Oxford: Oxford University Press.
50. Lorenz, K. 1981. "Dialogical Logic." In *Dictionary of Logic as Applied in the Study of Language*, edited by W. Marciszewski, 117–25. The Hague: Martinus Nijhoff.
51. Lorenzen, P., and K. Lorenz. 1978. *Dialogische Logik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
52. Majer, O., A.-V. Pietarinen, and T. Tulenheimo, eds. 2009. *Games: Unifying Logic, Language, and Philosophy*. Dordrecht: Springer.
53. Marion, M. 2006. "Hintikka on Wittgenstein: From Language-Games to Game Semantics." In *Truth and Games. Essays in Honour of Gabriel Sandu*, edited by T. Aho and A.-V. Pietarinen, 255–74. Helsinki: Societas Philosophicas Fennica.
54. Marion, M. 2009. "Why Play Logical Games?" In *Games: Unifying Logic, Language, and Philosophy*, edited by O. Majer, A.-V. Pietarinen, and T. Tulenheimo, 3–26. Dordrecht: Springer.
55. Marion, M. 2009. "Radical Anti-realism, Wittgenstein, and the Length of Proofs." *Synthese* 171:419–32.
56. Marion, M. 2010. "Between Saying and Doing: From Lorenzen to Brandom and Back." In *Construction. Festschrift for Gerhard Heinzmann*, edited by P. E. Bour, M. Rebuschi, and L. Rollet, 489–97. London: College Publications.
57. Marion, M., and M. Sadrzadeh. 2004. "Reasoning About Knowledge in Linear Logic: Modality & Complexity." In *Logic, Epistemology and the Unity of Science*, Cognitive Science Series, edited by D. Gabbay, S. Rahman, J. M. Torres, and J.-P. Van Bendegem, 327–50. Dordrecht: Hermes.
58. McDowell, J. 1981. "Anti-realism and the Epistemology of Understanding." In *Meaning and Understanding*, edited by H. Parret and J. Bouveresse, 225–48. Berlin: de Gruyter.
59. McGinn, C. 1980. "Truth and Use." In *Reference, Truth and Reality*, edited by M. Platts, 19–40. London: Routledge & Kegan Paul.
60. Pagin, P. 2004. "Is Assertion Social?" *Journal of Pragmatics* 36:833–59.
61. Pagin, P. 2009. "Assertion." In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. http://plato.stanford.edu/entries/assertion/.
62. Panaccio, C. 1999. *Le discours intérieur de Platon à Guillaume d'Ockham*. Paris: Éditions du Seuil.
63. Prawitz, D. 1977. "Meaning and Proofs: On the Conict Between Classical and Intuitionistic Logic." *Theoria* 43:1–40.
64. Priest, G. 1979. "Two Dogmas of Quineanism." *Philosophical Quarterly* 29:289–301.
65. Prior, A. N. 1960. "The Runabout Inference-Ticket." *Analysis* 21:38–39.
66. Quine, W. V. O 1986. *Philosophy of Logic* (2nd Edition). Cambridge MA: Harvard University Press.
67. Rahman, S., and L. Keiff. 2005. "How to Be a Dialogician." In *Logic, Thought and Action*, edited by D. Vanderveken, 359–408. Dordrecht: Springer.
68. Rahman, S., and T. Tulenheimo. 2009. "From Games to Dialogues and Back." In *Games: Unifying Logic, Language, and Philosophy*, edited by O. Majer, A.-V. Pietarinen, and T. Tulenheimo, 153–208. Dordrecht: Springer.
69. Ranta, A. 1988. "Propositions as Games as Types." *Synthese* 76:377–95.
70. Rebuschi, M. 2009. "Implicit Versus Explicit Knowledge in Dialogical Logic." In *Games: Unifying Logic, Language, and Philosophy*, edited by O. Majer, A.-V. Pietarinen, and T. Tulenheimo, 229–46. Dordrecht: Springer.
71. Restall, G. 2000. *An Introduction to Substructural Logics*. Oxford: Clarendon Press.
72. Robinson, R. 1953. *Plato's Earlier Dialectic*. Oxford: Clarendon Press.
73. Rödl, S. 2007. *Self-Consciousness*. Cambridge, MA: Harvard University Press.
74. Rückert, H. 2001. "Why Dialogical Logic?" In *Essays on Non-Classical Logic*, edited by H. Wansing, 165–85. Hackensack, NJ/London: World Scientific.

75. Rückert, H. 2007. "Dialogues as Dynamic Framework for Logic." Doctoral diss., University of Leiden. http://www.phil.uni-mannheim.de/fakul/phil2/rueckert/pdf/Rueckert_PhD_Dialogues.pdf.
76. Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson.
77. Scharp, K. 2003. "Communication and Content: Circumstances and Consequences of the Habermas-Brandom Debate." *International Journal of Philosophical Studies* 11/1:43–61.
78. Schmidt, D. A. 1997. "On the Need for a Popular Formal Semantics." *ACM SIGPLAN Notices* 32(1):115–16.
79. Schroeder-Heister, P., and K. Dosen, eds. 1993. *Substructural Logics*. Oxford: Clarendon Press.
80. Sellars, W. 1997. *Empiricism and the Philosophy of Mind*. Cambridge, MA: Harvard University Press.
81. Shapiro, S. 2000. "The Status of Logic." In *New Essays on the* a priori, edited by P. Boghossian and C. Peacocke, 333–66. Oxford: Clarendon Press.
82. Shieh, S. 1998. "On the Conceptual Foundations of Anti-realism." *Synthese* 115:33–70.
83. Shieh, S. 1998. "Undecidability in Anti-realism." *Philosophia Mathematica* 6:324–33.
84. Smiley, T. 1973. "What Is a Syllogism?" *Journal of Philosophical Logic* 2:136–54.
85. Stenius, E. 1967. "Mood and Language-Game." *Synthese* 17:254–74.
86. Sundholm. G. 1986. "Proof Theory and Meaning." In *Handbook of Philosophical* Logic, edited by D. Gabbay and F. Guenthner, vol. III, 471–506. Dordrecht: D. Reidel.
87. Tennant, N. 1979. "Language Games and Intuitionism." *Synthese* 42:297–314.
88. Tennant, N. 1987. *Anti-Realism and Logic. Truth as Eternal*. Oxford: Clarendon Press.
89. Watson, G. 2004. "Asserting and Promising." *Philosophical Studies* 117:57–77.
90. Wittgenstein, L. 1961. *Tractatus Logico-Philosophicus.* Translated by D. Pears and B. F. McGuinness. London: Routledge & Kegan Paul.
91. Wittgenstein, L. 2009. *Philosophical Investigations* (4th Edition). Translated by G. E. M. Anscombe, P. M. S. Hacker, and J. Schulte. Oxford: Blackwell.

Chapter 9
Conservativeness and Eliminability for Anti-Realistic Definitions

Towards a Global View of the Meaning of Logical Constants

Francesca Poggiolesi

9.1 Realistic Conservativeness and Eliminability

When it comes to the meaning of logical constants, we are far from having an unique conception of what this meaning consists of. There exist different semantic paradigms (e.g. see [12]). In this paper we only take into account two views on the meaning of logical constants: the realistic and anti-realistic ones.

According to a realistic conception, we grasp the meaning of a sentence when we know what it is for that sentence to be true, where the truth is thought of as something that a sentence either possesses or lacks independently from our capacity of recognising it. Let us consider a logical constant $\star$ of a language $\mathcal{L}$, and let us suppose that we want to define the meaning of $\star$ from a realistic point of view. This involves two steps: first, we specify a language $\mathcal{M}$ that does not contain $\star$, and then we formulate in $\mathcal{M}$ plus $\star$ an equivalence between sentence(s) containing $\star$ and the same sentence(s) containing other appropriate symbol(s) belonging to the language $\mathcal{M}$; by means of this equivalence we specify the truth values of the sentences containing $\star$. According to this explanation, the following equivalence:

$$\alpha \text{ is true and } \beta \text{ is true if, and only if, } \alpha \wedge \beta \text{ is true}$$

expressed in a language $\mathcal{M}$ plus $\wedge$ is nothing but the realistic definition of the symbol $\wedge$ of a language $\mathcal{L}$. This equivalence gives the meaning of the symbol $\wedge$ since it gives the truth-conditions under which the sentence $\alpha \wedge \beta$ is true. Note that what is on the right of the "if and only if" is called the *definiendum*, while what is on the left of the "if and only if" is called the *definiens*.

Once the explanation of a realistic definition is given, a question should naturally arise: what can ensure that a realistic definition gives us the whole meaning of the symbol that it defines and nothing more? The Polish logician S. Leśniewski [6] answered the question by introducing two criteria that are today well-known as *conservativeness* and *eliminability*.

F. Poggiolesi (✉)
CLWF, Vrije Universiteit Brussel, Brussels, Belgium
e-mail: poggiolesi@gmail.com

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_9,

CR {realistic conservativeness} there exists no formula α, not containing any occurrence of the symbol $\star$, that is valid in $\mathcal{L}$ without already being valid in $\mathcal{L}$ minus $\star$.

ER {realistic eliminability} for any formula α of $\mathcal{L}$ containing the symbol $\star$, theres exists a formula β not containing the symbol $\star$, such that β is semantically equivalent to α.

These two criteria ensure that a definition gives the meaning of the symbol it defines and nothing more. More precisely, conservativeness ensures that the definition gives nothing more than the meaning, while eliminability ensures that it gives the whole meaning.

9.2 Anti-Realistic Definitions and Sequent Calculus

The ontological conception of the meaning provided by realism was substituted by anti-realism with an epistemic conception. Indeed, according to an anti-realistic conception, the meaning of an expression is determined by its use. This idea can be found in Wittgenstein, who claims that

> [f]or a *large* class of cases—though not for all—in which we employ the world 'meaning,' it can be defined thus: the meaning of a word is its use in the language. [13, §43]

So the meaning of a sentence no longer depends on the existence of properties such as truth or falsity, but on our capacity of employing that sentence.

What about the meaning of logical constants? Even in this case the meaning is determined by the use. What kind of use do we make of logical constants? A purely inferential one. The conclusion that we can draw is then that the meaning of logical constants is fully defined by the inferential rules that govern their use. Indeed this is the position frequently labelled as *inferentialism*.

Traditionally, the inferential rules that inferentialism refers to are those of natural deduction calculi. As suggested by Paoli [7], this is due to a twofold reason: on the one hand, the rules of natural deduction calculi have a greater intuitive appeal, on the other hand, the natural deduction calculi themselves are better suited to handle intuitionistic logic, which in turn was the logic preferred by authors supporting anti-realistic positions. However, more recently (see [3, 7, 12]), the sequent calculus has been rediscovered by the inferentialism. This is again based on two grounds: first, intuitionistic logic is no longer the only constructive logic; second, in the sequent calculus, the clear separation between logical and structural rules helps understanding the roles that the logical rules of the sequent calculus play in the definition of logical constants.

In the rest of the paper, we follow this new trend and we assume that the left and right introduction rules of the sequent calculus give the meaning of the constant they introduce. Indeed, the left and right introduction rules of the sequent calculus are taken to provide the grounds for inferring a sentence containing the connective that they define, in the antecedent and in the consequent of a sequent, respectively.

Let us end the current section by illustrating the way we formulate an anti-realistic definition of a constant $\star$ of a language $\mathcal{L}$. First, a sequent calculus T that does not contain $\star$ must be specified; then, the constant $\star$ must be added to the sequent calculus T via a set of left and right introduction rules. According to this explanation, the following rules[1]:

$$\frac{\alpha, \beta, M \Rightarrow N}{\alpha \wedge \beta, M \Rightarrow N} \wedge L^* \qquad \frac{M \Rightarrow N, \alpha \qquad M \Rightarrow N, \beta}{M \Rightarrow N, \alpha \wedge \beta} \wedge R$$

added to a calculus T, are nothing but an anti-realistic definition of the symbol $\wedge$. These rules give the meaning of the symbol $\wedge$ since they give the grounds under which the sentence $\alpha \wedge \beta$ can be asserted on the right and on the left of the sequent.

To sum up the present discussion, we will say that in the realistic case the central notions are those of language and truth, and that definitions take the form of equivalencies; in the anti-realistic case, on the other hand, the central notions are those of calculus and proof and definitions take the form of rules.

Realistic Definitions	Language	Truth	Equivalence
Anti-realistic Definitions	Calculus	Proof	Rules

9.3 Anti-Realistic Conservativeness

Logical rules are (anti-realistic) definitions of the symbol they introduce. In order to be sure that they are good definitions, we should check that they give the whole meaning of the symbol they introduce and nothing more. This amounts to verify that they satisfy the conservativeness and eliminability criteria. In the first section we have introduced these criteria for realistic definitions. In the current and in the next section we will attempt at reformulate the conservativeness and eliminability criteria in order to adapt them to the anti-realistic case.

We start from the conservativeness criterion. This criterion ensures that a definition does not give anything more than the meaning of the symbol that it defines. In the realistic case, meaning is expressed in terms of truth values. Hence realistic conservativeness requires that a definition does not modify the truth value of the sentences not containing the symbol to define. Indeed, if the definition affected the truth-value of the sentences not containing the symbol to define, then it would give more than the pure meaning of the symbol to define, i.e. it would say something about the meaning, expressed by truth-values, of other symbols.

[1] α, β stand for formulas, while M, N stand for multisets of formulas. The definition of sequent is the standard one.

Let us then apply an analogous reasoning to anti-realistic definitions. In the anti-realistic case meaning is expressed in terms of the conditions under which an expression can be asserted. Hence, anti-realistic conservativeness must require that a definition does not modify the asseribility, i.e. the provability, of sentences non containing the symbol to define. More precisely we have:

CA {anti-realistic conservativeness} a calculus $T^{'}$ obtained by adding to the calculus T one or more connectives and rules concerning these connectives, is said to conservatively extend the calculus T, when $T^{'}$ proves no sequent containing just the old connectives which was not already provable in T.

We thus have anti-realistic conservativeness. Now we need to find a property of the sequent calculus that, if satisfied, guarantees that the logical rules are conservative. According to many (e.g. see [69, 92, 144]), such a property is the subformula property. Indeed, if a sequent calculus $T^{'}$ enjoys the subformula property, any formula that occurs in a proof of T must occur as a subformula of the end-sequent itself, and this is enough to ensure that any new rule of $T^{'}$ needs not to be used for establishing this end-sequent.

9.4 Anti-Realistic Eliminability

We have introduced anti-realistic conservativeness. Now it is the turn of anti-realistic eliminability. In order to define anti-realistic eliminability, we are going to proceed as in the last section: first we analyse eliminability for the realistic case, and then we exploit such an analysis for reformulating eliminability in a way that suites the anti-realist case.

Eliminability is a criterion that ensures that a definition gives the whole meaning of the symbol it defines. In the realistic case a definition is an equivalence, and meaning is expressed in terms of truth values. Therefore, realistic eliminability requires that, for any sentence containing the symbol to define, there is a sentence not containing the symbol to define, that has the same meaning, i.e. the same truth value.

Let us try to apply an analogous reasoning to anti-realistic definitions. In the anti-realistic case a definition is a set of rules and meaning is expressed in terms of the conditions under which an expression can be asserted. Therefore it seems reasonable to think that anti-realistic eliminability should require that:

EAI {anti-realistic eliminability} the set of rules exactly determine which sentences containing the new symbol can be asserted, i.e can be proved.

We thus have anti-realistic eliminability. Our task is still not finished. We have to find a property of the sequent calculus that ensures that the logical rules satisfy the eliminability criterion. The solution to this problem is not too difficult. Indeed the proof of the admissibility of the structural rules is exactly what we need. If this requirement is not met, then there will be sequents involving the new constant which cannot be established solely on the basis of the logical rules for that symbol together with the old rules, restricted to the old vocabulary. If this requirement is met, on the other hand, then the logical rules for the new symbol enable us to infer all the sequents involving the new vocabulary, without having to apply the old rules except for formulas involving the old vocabulary.[2]

9.5 Logical Variant of the Sequent Calculus

Logical rules of the sequent calculus are anti-realistic definitions. In order to be definitions that give the whole meaning of the symbol they define and nothing more, they should satisfy the conservativeness and eliminability criteria. We now know that they satisfy these criteria, if the sequent calculus that they belong to, have two properties: on the one hand, it must satisfy the subformula property,[3] and on the other hand the structural rules must be admissible in it. The goodness of the rules seems therefore to depend on the context that these rules belong to. Note that this conclusion is, as far as we know, quite original with respect to the current literature on the topic. Indeed, although inferentialists (e.g. see [12]) have proposed many properties that logical rules should satisfy in order to be considered as "good" definitions, these properties always have a local flavour. The separation property, for example, requires that logical rules should not introduce more than one constant a time. The conclusion that we have drawn here is, on the contrary, that the logical rules of the sequent calculus not only should satisfy the standard local constraints, but also, and perhaps most importantly, context constraints.

Let us now illustrate within an example our conclusions. For this purpose, let us consider the logical constant $\wedge$ and the sequent calculus of classical logic settled up in the following way. **Gcl** is composed of:

Axioms

Ax: $p, M \Rightarrow N, p$ $\qquad\qquad$ $\bot A$: $\bot, M \Rightarrow N$

Propositional Rules

$$\frac{\alpha, \beta, M \Rightarrow N}{\alpha \wedge \beta, M \Rightarrow N}\wedge L^* \qquad\qquad \frac{M \Rightarrow N, \alpha \qquad M \Rightarrow N, \beta}{M \Rightarrow N, \alpha \wedge \beta}\wedge R$$

[2] Note that [5] reaches an analogous conclusion.

[3] A calculus satisfies the subformula property when: (i) it is cut-free, and (ii) in each of its rules, the formulas that belong to the premises are subfomulas of the formulas that belong to the conclusions.

$$\frac{\alpha, M \Rightarrow N \quad \beta, M \Rightarrow N}{\alpha \vee \beta, M \Rightarrow N} \vee L \qquad \frac{M \Rightarrow N, \alpha, \beta}{M \Rightarrow N, \alpha \vee \beta} \vee R^*$$

$$\frac{M \Rightarrow N, \alpha \quad \beta, M \Rightarrow N}{\alpha \rightarrow \beta, M \Rightarrow N} \rightarrow L \qquad \frac{\alpha, M \Rightarrow N, \beta}{M \Rightarrow N, \alpha \rightarrow \beta} \rightarrow R^*$$

Gcl is a variant of the classical sequent calculus that has been introduced by Dragalin [2] and that is called *logical* variant since, as the reader can easily see, it does not contain any structural rule. **Gcl** enjoys the subformula property since the cut-rule is admissible and in its rules the formulas contained in the premises are subformulas of the formulas contained in the conclusion. Moreover, the structural rules of contraction and weakening are admissible. Therefore, in **Gcl**, the rules $\wedge L^*$ and $\wedge R$ satisfy the conservativeness and eliminability criteria and hence they give the whole meaning of the symbol $\wedge$ and nothing more. We can then conclude that, in the case of classical logic, **Gcl**[4] represents the context we should look at, if we want to have good definitions of logical constants.

More generally, we can stress the importance for a logic to have a (logical) variant of the sequent calculus analogous to **Gcl**. Indeed, in this variant of the sequent calculus, the structural rules can be shown to be admissible and therefore the logical rules satisfy the two criteria of conservativeness and eliminability. In other words, in this variant of the sequent calculus the logical rules can be taken as good definitions of the symbol that they define.

9.6 The Modal Case

Let us now come to the case of modal logic. Modal logic is the logic obtained by adding to classical logic the symbol $\Box$.[5] This symbol can be interpreted in many different ways, such as "it is necessary that," "it is known that," "it is compulsory that." According to the different interpretations, different systems of modal logic exist. The most well-known are: **K**, **KT**, **S4**, **S5**. In what follows, we will concentrate on the Hilbert system **S5** where the constant $\Box$ is meant to represent the concept of necessity.

Let us remind the reader that the Hilbert system **S5** is composed of:

– classical propositional logic
– distribution axiom: $\Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)$
– axiom T: $\Box\alpha \rightarrow \alpha$; axiom 4: $\Box\alpha \rightarrow \Box\Box\alpha$; axiom B: $\alpha \rightarrow \Box\neg\Box\neg\alpha$
– modus ponens and rule of necessitation (from α, derive $\Box\alpha$)

---

[4] **Gcl** as well as all the other logical variants of the sequent calculus.

[5] We assume as primitive the only modal constant $\Box$. The constant $\Diamond$ can be obtained from $\Box$ by the standard definition: $\Diamond\alpha := \neg\Box\neg\alpha$

Let us suppose that we want to have a realistic definition of the constant $\Box$ of the system **S5**. In this case no problem arises thanks to well-known Kripke semantics equivalence:

$$i \models \Box\alpha \text{ if, and only if, } \forall j(iRj \rightarrow j \models \alpha)$$

where the relation R has the following three properties: reflexivity, symmetry and transitivity.

Let us on the contrary suppose that we want to have an anti-realistic definition of the constant $\Box$ of the system **S5**. In this case there is a little consensus on the logical rules that are to be taken as definitions of $\Box$. The disagreement is due to the fact that the system **S5** seems to be particularly difficult to capture in a Gentzen framework. According to our conclusions, the search for rules defining the modal constant $\Box$ in **S5** is even more problematic, since we do not need a whatever sequent calculus for **S5**, but one that enjoys the subformula property and also where the structural rules are all provable to be admissible. (I.e. we need a logical variant of the sequent calculus for the system **S5**.)

In a recent article [8], Poggiolesi has introduced an hypersequent calculus for the system **S5** that seems to have all the characteristics required for our purpose. In order to introduce this calculus, let us, first of all, give the following two definitions.

Definition 1 A hypersequent is a syntactic object of the form:

$$M_1 \Rightarrow N_1 | M_2 \Rightarrow N_2 | \ldots | M_n \Rightarrow N_n$$

where $M_i \Rightarrow N_i$ $(i = 1, \ldots, n)$ are classical sequents. G, H, ... will denote hypersequents.

Definition 2 The intended interpretation of a hypersequent is defined inductively in the following way:

- $(M \Rightarrow N)^\tau := \bigwedge M \rightarrow \bigvee N$,
- $(\Gamma_1 | \Gamma_2 | \ldots | \Gamma_n)^\tau := \Box\Gamma_1^\tau \vee \Box\Gamma_2^\tau \vee \ldots \vee \Box\Gamma_n^\tau$

The postulates of the calculus **Gs5** are:

Initial Hypersequents

$$G|\ p, M \Rightarrow N, p \qquad\qquad G|\bot, M \Rightarrow N$$

Propositional Rules

$$\frac{G|\alpha, \beta, M \Rightarrow N}{G|\alpha \wedge \beta, M \Rightarrow N}\ {\wedge L^*} \qquad\qquad \frac{G|M \Rightarrow N, \alpha \qquad G|M \Rightarrow N, \beta}{G|M \Rightarrow N, \alpha \wedge \beta}\ {\wedge R}$$

$$\frac{G|\alpha, M \Rightarrow N \qquad G|\beta, M \Rightarrow N}{G|\alpha \vee \beta, M \Rightarrow N} \vee L \qquad \frac{G|M \Rightarrow N, \alpha, \beta}{G|M \Rightarrow N, \alpha \vee \beta} \vee R^*$$

$$\frac{G|M \Rightarrow N, \alpha \qquad G|\beta, M \Rightarrow N}{G|\alpha \rightarrow \beta, M \Rightarrow N} \rightarrow L \qquad \frac{G|\alpha, M \Rightarrow N, \beta}{G|M \Rightarrow N, \alpha \rightarrow \beta} \rightarrow R^*$$

Modal Rules

$$\frac{G|\alpha, \Box\alpha, M \Rightarrow N}{G|\Box\alpha, M \Rightarrow N} \Box L_1 \qquad \frac{G|M \Rightarrow N| \Rightarrow \alpha}{G|M \Rightarrow N, \Box\alpha} \Box R$$

$$\frac{G|\Box\alpha, M \Rightarrow N|\alpha, P \Rightarrow Q}{G|\Box\alpha, M \Rightarrow N|P \Rightarrow Q} \Box L_2$$

Let us try to abstract from the fact that we are dealing with hypersequents (and that therefore the rules might appear a little bit more complicated than what they usually are), and let us concentrate on the sequents that are displayed in each rule of the calculus. The propositional rules are just the classical ones; to realise it, it suffices to compare them with those that were introduced in Section 9.5 and that composed the calculus **Gcl**. Let us then move to analyse the modal rules. Let us first observe that in the case of the modal rules, differently from the case of the propositional rules, we do not have just one sequent displayed but two. This just means that in the case of modal rules we need more structure: a second sequent and the meta-linguistic symbol slash. Let us now concentrate on the conditions of asseribility of the constant $\Box$ that the modal rules provide. The rule $\Box R$ tells us that the grounds for asserting the formula $\Box\alpha$ on the right side of the sequent $M \Rightarrow N$ consists in having a sequent different from $M \Rightarrow N$ that only contains α on its right side. The rules $\Box L_1$ and $\Box L_2$, considered together, tell us, on the contrary, that the grounds for inferring the formula α on the left side of the sequent $M \Rightarrow N$ consists in having the formula α in whatever sequent, $M \Rightarrow N$ included, that compose the hypersequent G.

Let us now examine the calculus **Gs5**. We can easily remark that the calculus **Gs5** enjoys the following three properties. First of all, in each rule of the calculus, the formulas that occur in the premises are subformulas of the formulas that occur in the conclusion. Secondly, the cut-rule is shown to be admissible (see [8]). Thirdly, any other structural rule is shown to be admissible (see [8]).

The conclusion that we can draw is that the conservativeness and eliminability criteria are satisfied by the three modal rules and that therefore these rules give the whole meaning of the symbol $\Box$ and nothing more. These rules are thus the anti-realistic definitions of the symbol $\Box$ in **S5**, i.e. these rules are the definitions of (the formalisation of) the concept of necessity.

9.7 Anti-Realistic Definitions in Past Attempts

In this paper our starting point were certain observations made by Lésniewski in 1929 concerning definitions. Following Lésniewski, if definitions are supposed to give the meaning of the symbol that they define, then they should be conservative and eliminable, since these two criteria ensure that a definition gives the whole meaning of the symbol to define and nothing more. The two criteria of conservativeness and eliminability proposed by Lésniewski are nowadays commonly accepted. However, it should be stressed that Lésniewski formulated these two criteria in order to fit realistic definitions: indeed, he employed the notions of language and of truth. But today realistic definitions are no longer the only definitions on the market. Most notably we also have anti-realistic definitions. Of course even this alternative type of definitions should be conservative and eliminable. The firsts who tried to defend this position were Belnap [1] and Hacking [3]. Nevertheless, in their attempt, they neglected two important features: first, they did not take into account the eliminability criterion, and secondly, they did not reformulate Lésniewski's definitions of conservativeness in order to fit with the anti-realistic case. Now it is clear that, if this reformulation is not operated, severe conceptual problems might arise, since the concepts that are at the basis of anti-realistic definitions are completely different from those that support realistic definitions. In this paper, not only we did not neglect the eliminability criterion, but moreover we reformulated the conservativeness criterion in a way that suits anti-realistic definitions.

Our approach to the notion of conservativeness not only differs from that of Belnap and Hacking, but also from that commonly adopted by inferentialists. One of the main attacks against inferentialism was famously given by Prior [9] who introduced the following two tonk-rules

$$\frac{\alpha}{\alpha\ tonk\ \beta} \qquad \frac{\alpha\ tonk\ \beta}{\beta}$$

in a natural deduction system with the aim of showing that inferentialism was ill-founded. Indeed, the calculus resulting from their addition can prove $\alpha \vdash \beta$ for any formula α and β whatsoever, which is of course to say that it is trivial.

The consequence that Prior drew is that not any set of rules is meaning conferring. Although many commentators have taken this as an argument against inferentialism, defenders of this position replied that there are natural constraints on logical rules which guarantee them to confer meaning to the constant they introduce, and that tonk and tonkish connectives do not satisfy these constraints. The constraint that has received by far the most attention is proof-theoretic *harmony*.

Informally speaking, harmony is supposed to balance two features of a logical connective $\star$: (i) the conditions under which one is entitled to assert a sentence containing the $\star$ connective; (ii) the consequences one is entitled to draw from a sentence containing the $\star$ connective. There have been many attempts to formally capture the harmony requirement. If inferentialists have ever considered the conservativeness criterion, it was precisely as a possible formalisation of the harmony requirement.

Given these remarks, the originality of our approach with respect to inferentialism is then evident. In this paper, we have never claimed that logical rules should be conservatives for the harmony's sake, but because they are (anti-realistic) definitions, and definitions in general should satisfy this criterion. (Note that this way we are also saved from the criticisms (e.g. see [10]) against conservativeness as a formalisation of the harmony criterion.)

The main consequence that our investigation had led us to, is the fact that, in order to understand the meaning of logical constants, we cannot prevent ourselves at considering the sequent calculus that these rules belong to. This position seems to contradict what Paoli [7] supports in a recent article. Indeed, following Paoli, we can make a distinction between the *operational* and the *global* meaning of logical constants: while the operational meaning of a logical constant $\star$ is fully specified by the right and left introduction rules for $\star$, the global meaning of a logical constant $\star$ is fully specified by the class of provable sequents containing $\star$. The two notions come apart: while we can choose to always deal with the same left and right introduction rules even when treating with different logics, the set of provable sequents obviously varies from logic to logic. Our position in this paper have strong relations with the idea of a global meaning, but obviously contrasts with the idea of an operational meaning. Unluckily, Paoli suggests identifying the meaning of logical constants with the operational meaning. We thus feel willing to defend our view.

A part from the fact that Paoli does not give any argument neither in defence of the distinction between operational and global meaning, nor in support of the identification of meaning with operational meaning, we think that there are at least two arguments against Paoli's operational meaning. The first one has been proposed by Hjortland [4, p. 16]:

> Inferentialism leaves it open whether all inferential rules are meaning-conferring or only some (and does not even consider structural assumptions not in rule-form), but meaning-theoretically the choice makes considerable difference. [...] Inferentialism is based on the idea that, at least for logical constants, the entrenched use of the expressions fully determines their meaning. But, if some aspects of the inferential role of these expressions come short of being semantically significant, then we need a corresponding use-theoretic distinction to explain how meaning supervenes on some (systematic) use of an expression but not all (systematic) use of an expression.

In order to explain the second argument, we have to introduce an important distinction. It is a well-known fact that in the sequent calculus for classical logic, for each of the three connectives $\wedge$, $\vee$, and $\rightarrow$, we can choose between two equivalent formulations of the left introduction rules and two equivalent formulations of the right introduction rules. As a way of example, let us illustrate the four different rules that introduce the constant $\wedge$:

$$\frac{\alpha_i, M \Rightarrow N}{\alpha_0 \wedge \alpha_1, M \Rightarrow N}\wedge L \qquad \frac{M \Rightarrow N, \alpha \quad M \Rightarrow N, \beta}{M \Rightarrow N, \alpha \wedge \beta}\wedge R$$

$$\frac{\alpha, \beta, M \Rightarrow N}{\alpha \wedge \beta, M \Rightarrow N} \wedge L^* \qquad \frac{M \Rightarrow N, \alpha \quad P \Rightarrow Q, \beta}{M, P \Rightarrow N, Q, \alpha \wedge \beta} \wedge R^*$$

The rules where the context does not change—in the case of conjunction, the rules $\wedge L$ and $\wedge R$—are called *additive*. The rules where the context varies—in the case of conjunction, the rules $\wedge L^*$ and $\wedge R^*$—are called *multiplicative*. The equivalence of additive and multiplicative rules is provable by means of the structural rules of weakening and contraction (for further details see [11]).

Let us now move back to the second argument against Paoli's position. This argument is developed in the following way. If the rules of inference *tout court* determine the meaning of the constants they introduce, i.e. if the meaning of the constants is their operational meaning, then we must accept that in the classical sequent calculus the constants $\wedge$, $\vee$ and $\rightarrow$ may have two different meanings: an additive and a multiplicative one. This clearly sounds as an unreasonable conclusion. To counter it, one could reply that the above conclusion is not entirely correct since additive and multiplicative rules can be shown to be equivalent, and so that the different meanings they provide are in reality the same. This is certainly true, but crucially the equivalence between the additive and multiplicative rules can be shown to hold *only* from a global point of view since we require structural rules to prove it. Hence the reply is not decisive and also the second argument would appear to go against the operational meaning.

Once the operational meaning has been ruled out, the only choice that rests is the global meaning. Shall we accept the global meaning just because we do not have any other option? In this paper we have tried to answer negatively to this question: there exist deep and important reasons in support of—if not exactly the global meaning suggested by Paoli—a global view of the meaning of logical constants.

References

1. Belnap, N. D. 1962. "Tonk, Plonk and Plink." *Analysis* 22:130–34.
2. Dragalin, A. G. 1988. *Mathematical Intuitionism. Introduction to Proof Theory*. Providence, RI: American Mathematical Society.
3. Hacking, I. 1979. "What Is Logic?" *Journal of Philosophy* 76:285–319.
4. Hjortland, O. 2007. "Proof-Theoretic Harmony and Structural Assumptions", 1–20. http://olethhjortland.googlepages.com/papers.
5. Kremer, M. 1988. "Logic and Meaning: The Philosophical Significance of the Sequent Calculus." *Mind, New Series* 97:50–72.
6. Leśniewski, S. 1929. "Gründzuge eines neuen systems der grundlagen der mathematik." *Fundamenta Mathematicae* 14:1–81.
7. Paoli, F. 2003. "Quine and Slater on Paraconsistency and Deviance." *Journal of Philosophical Logic* 32:531–48.
8. Poggiolesi, F. 2008. "A Cut-Free Simple Sequent Calculus for Modal Logic S5." *Review of Symbolic Logic* 1:3–15.
9. Prior, A. N. 1960. "A Runabout Inference Ticket." *Analysis* 21:38–39.

10. Read, S. 2000. "Harmony and Autonomy in Classical Logic." *Journal of Philosophical Logic* 29:123–54.
11. Troelestra, A. S., and H. Schwichtenberg. 1996. *Basic Proof Theory*. Cambridge, MA: Cambridge University Press.
12. Wansing, H. 2000. "The Idea of a Proof-Theoretic Semantics and the Meaning of the Logical Operations." *Studia Logica* 64:3–20.
13. Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.

Chapter 10
Realism, Antirealism, and Paraconsistency

Graham Priest

10.1 Introduction

The debate between realists and antirealists (about various topics) has occasioned an enormous literature in the last 35 years.[1] Usually this is carried out in terms of the contrast between classical and intuitionist logic. Intuitionist logic is not, of course, the only non-classical logic.[2] Another important class of such logics comprises paraconsistent logics. How do they fit into the debate? This note answers the question. Paraconsistency, as such, is neutral to the debate, in the sense that there are paraconsistent logics that are as unfriendly to antirealism as classical logic; and there are paraconsistent logics that are as susceptible to an antirealist understanding as intuitionist logic. I will show this by considering just one family of paraconsistent logics: those that have a binary relational semantics.

10.2 Classical vs. Intuitionist Logic

The heart of the realist/antirealist debate about some matter concerns whether sentences about it are to be given truth conditions where the notion of truth in question is verification-transcendent. Thus, consider classical propositional logic. We may suppose that the language contains the connectives $\vee$, $\wedge$, $\supset$, and $\neg$. An evaluation is a function, ν, which assigns a truth value (1 or 0) to every propositional parameter. This is then extended to such a map for all formulas by the familiar conditions. For all formulas, A and B:

$$\nu(A \vee B) = 1 \text{ iff } \nu(A) = 1 \text{ or } \nu(B) = 1$$

[1] For a gentle introduction, see [11, chap. 8].

[2] For an introduction to non-classical logics, see Priest [6, 9].

G. Priest (✉)
Departments of Philosophy, University of Melbourne, Melbourne, VIC, Australia

University of St Andrews, St Andrews, UK

The Graduate Center, City University of New York, New York, NY, USA
e-mail: g.priest@unimelb.edu.au

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_10, © Springer Science+Business Media B.V. 2012

$\nu(A \wedge B) = 1$ iff $\nu(A) = 1$ and $\nu(B) = 1$
$\nu(A \supset B) = 1$ iff $\nu(A) = 0$ or $\nu(B) = 1$
$\nu(\neg A) = 1$ iff $\nu(A) = 0$

Validly is defined as terms of truth preservation in all evaluations.

The truth conditions for negation are such that $\neg A$ is true simply if A fails to be true. Generally speaking, we may have no way of telling of a sentence that it fails to be true. We would appear to have no way of knowing, for example, whether 'There are unicorn-like creatures at some place and time in the cosmos', fails to be true. The truth condition of negation are therefore verification-transcendent. More generally, the logic verifies the Law of Excluded Middle (LEM): for all A, $A \vee \neg A$ is valid. Yet for some As we may have no way of verifying either A or $\neg A$.

Compare this with intuitionist logic. This can be given a familiar Kripke-style semantics.[3] An interpretation is a structure $\langle W, R, \nu\rangle$. W is a non-empty set of worlds. These are to be thought of as states of information. Each contains the things that have been verified at a certain stage. That is:

$\nu_w(A) = 1$ iff A is verified at w

R is a binary accessibility relation on the worlds. $w_1 R w_2$ means that w_2 is a possible state of information obtained from w_1 by adding some number (possibly zero) of verifications. It is therefore reflexive and transitive. ν is a map which assigns a truth value, $\nu_w(p)$ (1 or 0), to each propositional parameter, p, at each world, w. There is a heredity constraint:

if $\nu_{w_1}(p)$ and $w_1 R w_2$ then $\nu_{w_2}(p)$

What is verified stays verified. (Given the truth conditions for the connectives, this extends to all formulas.) The truth conditions for the connectives are given as follows:

$\nu_w(A \vee B) = 1$ iff $\nu_w(A) = 1$ or $\nu_w(B) = 1$
$\nu_w(A \wedge B) = 1$ iff $\nu_w(A) = 1$ and $\nu_w(B) = 1$
$\nu_w(A \supset B) = 1$ iff for all w' such that wRw', $\nu_{w'}(A) \neq 1$ or $\nu_{w'}(B) = 1$
$\nu_w(\neg A) = 1$ iff for all w' such that wRw', $\nu_{w'}(A) = 0$

Validity is defined in terms of truth preservation at all worlds of all interpretations.

The truth conditions for negation say that $\neg A$ holds at a world if at no further worlds A holds. Intuitively, the only way for this to happen is for us to have a verification that A will never be verified. (If we have such a verification, A will never be verified. Conversely, if we have no such verification, then there is a possible

[3] See [6, §6.3].

future in which A is verified.) Hence, the semantics give a plausible representation of the fact that the truth conditions of negation may be understood in terms of verification.

Similarly, the truth conditions of the other connectives can be thought of in terms of verification. A conjunction is verified iff both conjuncts are. A disjunction is verified iff a disjunct is. The truth conditions for the conditional say that we will never have A without B. The only way for this to happen is for us to have a construction that turns verifications of A into verifications of B. (If we have such a construction, any verification of A may be turned straightforwardly into a verification of B. Conversely, if we have no such construction, then there is a possible future in which A is verified but not B.)

10.3 The Logic of Constructible Negation

Let us now consider the logic of constructible negation, N_3, first proposed by Nelson [3]. This is essentially intuitionist logic with a different account of negation. It can be given a Kripke-style semantics as follows. An interpretation is a structure $\langle W, R, \rho\rangle$. W and R are as for intuitionist logic. ρ is a world-indexed relation between propositional parameters and $\{1, 0\}$, subject to the constraint that for every propositional parameter, p, and world, w:

Exclusion: it is not the case that $p\rho_w 1$ and $p\rho_w 0$

Nothing is both true and false. (And given the truth/falsity conditions for the connectives, this extends to all formulas.) Since truth and falsity are not independent, two heredity conditions are required. For every propositional parameter, p:

if $p\rho_{w_1} 1$ and $w_1 R w_2$ then $p\rho_{w_2} 1$
if $p\rho_{w_1} 0$ and $w_1 R w_2$ then $p\rho_{w_2} 0$

Again, given the truth conditions for the connectives, this extends to all formulas. The truth/falsity conditions for the connective are as follows:

$(A \vee B)\rho_w 1$ iff $A\rho_w 1$ or $B\rho_w 1$
$(A \vee B)\rho_w 0$ iff $A\rho_w 0$ and $B\rho_w 0$

$(A \wedge B)\rho_w 1$ iff $A\rho_w 1$ and $B\rho_w 1$
$(A \wedge B)\rho_w 0$ iff $A\rho_w 0$ or $B\rho_w 0$

$(A \supset B)\rho_w 1$ iff for all w' such that wRw', it is not the case that $A\rho_{w'} 1$ or $B\rho_{w'} 1$
$(A \supset B)\rho_w 0$ iff $A\rho_w 1$ and $B\rho_w 0$

$\neg A\rho_w 1$ iff $A\rho_w 0$
$\neg A\rho_w 0$ iff $A\rho_w 1$

Validity is defined, as usual, in terms of truth preservation at all worlds of all interpretations.

As far as the positive logic goes, this is just the same as intuitionist logic. The semantics vary only in their handling of negation.[4] The question is whether this makes good antirealist sense. In particular, what are we to make of the fact that $A\rho_w 0$?

The obvious answer is that, just as $A\rho_w 1$ means that A is verified at w, so $A\rho_w 0$ means that A ia falsified. Does the antirealist have an independent notion of falsification? The answer is 'yes'. There are ways of showing that something is false directly. Perception can be such a method. One can see directly that something is *not* the case. For example, when you enter a room you can see that Pierre is not there. You do not have to see that the things in the room are a table, a chair, ... and reason: Pierre is not a table, Pierre is not a chair, ... therefore Pierre is not in the room.[5] Moreover, if one is not tied to intuitonist logic, one can give a *direct* proof of a negated sentence. And given the fact that verification *and* falsification make perfectly good independent sense, then the logic of constructible negation makes perfectly good antirealist sense. Modulo this change, we tell exactly the same story as for intuitionist logic.

Given the understanding of direct falsification, the falsity conditions for the connectives are straightforward, with the exception of those for the conditional. Generally speaking, it is less than clear when one should take a conditional to be false. The Nelson conditions say that a conditional, $A \supset B$, is false iff there is an actual counter-example, $A \wedge \neg B$. It is certainly plausible that this is a sufficient condition. It is less clear that this is necessary.

Unlike intuitionist logic, N_3—and all the logics in the same family that we will go on to note—verifies both halves of the Law of Double Negation, and especially $\neg\neg A \models A$. From the antirealist perspective that N_3 provides, the failure of this in intuitionist logic is entirely an artifact of its semantic one-sidedness. Contraposition, on the other hand, fails: $A \to B \not\models \neg B \to \neg A$. Truth preservation forward does not guarantee falsity preservation backwards.

Note that it is possible to have a second negation in the language, which behaves as does intuitionist negation. Thus, suppose that there is a constant, $\perp$, that is false everywhere: an intuitionist-like negation may be defined as $A \supset \perp$. Then $A \supset \perp$ is true at w if A is true at no accessible world, and $A \supset \perp$ is false at w if A is true at w. We may think of $\neg A$ as expressing a direct falsification, and $A \supset \perp$ as expressing an inferential one. $\neg A$ is stronger than $A \supset \perp$. If $\neg A$ holds at w, it holds at all accessibile worlds, by heredity. Hence A fails at all accessible worlds, that is, $A \supset \perp$ is true at w: $\neg A \vDash A \supset \perp$. But it is quite possible to have $A \supset \perp$ true at

[4] On all this, see [9, §9.7a].

[5] This and other examples are discussed in [8, chap. 3] (esp. 3.5).

w, that is, A true at no accessible world, without having $\neg A$ true at w ($\neg A$ may, in fact, fail at all accessible worlds): $A \supset \bot \not\models \neg A$.[6]

10.4 Paraconsistency

N_3 is not a paraconsistent logic. In it, contradictions still entail everything. However, we get a paraconsistent logic, N_4, simply by dropping the Exclusion constraint.[7] In this, contradictions do not imply everything. Can we make sense of this liberalisation from an antirealist perspective?

In N_4, truth and falsity at a world, that is, verification and falsification, have *complete* independence. That one may be in a position to verify neither A nor $\neg A$, that is, in a position neither to verify nor to falsify A, is standard antirealist fare. The thought that one might be in a position to verify both A and $\neg A$, that is, in a position both to verify and to falsify A, is more radical. Yet it makes perfectly good sense. In some paradoxes, such as Berry's, for example, one can give a verification (direct proof) of some claim and a falsification (direct proof of the negation) of it. For another example, there are many terms in science that are multi-criterial; that is, for which we have more than one criterion for applying them. Obvious examples are temperature terms. That the fluid in some beaker has a temperature of 20°C can be verified by both a correctly functioning mercury thermometer and a correctly functioning electro-chemical thermometer. However, there is no reason a priori why these two criteria should hang together. It could be the case that, by one criterion, the fluid has a temperature of 20°C; yet by the other it has a temperature of 21°C, and so it does not have a temperature of 20°C. Arguably, there are places in the history of science where exactly this divergence of criteria has happened.[8] N_4, therefore, is a perfectly acceptable antirealist logic.[9]

These considerations show, incidentally, not only that there are paraconsistent logics that have antirealist interpretations, but that dialetheism is also quite compatible with antirealism. In the situation explained, A is both verified and falsified; that is, both A and $\neg A$ are true.

There is another logic with constructible negation in the vicinity of N_3 and N_4. Start with N_4, and augment this with the constraint which is the dual of Exclusion. For every propositional parameter, p, and world, w:

Exhaustion: either $p\rho_w 1$ or $p\rho_w 0$

[6] In N_4, which we shall meet in the next section, the inference in both directions fails. The fact that $\neg A$ is true at a world does not entail that A is not true there.

[7] N_4 was first proposed by Almukdad and Nelson [1]. N_3 and N_4 are discussed in [13], and also in [9, §9.7a], where they are called L_3 and L_4.

[8] The point is made in [4, chap. 1] and further discussed in [10, §2.II.i].

[9] Rumfitt [12] argues for treating truth and falsity even-handedly, in the way required by N_3 and N_4. He does so by analysing falsity in terms of a primitive notion of rejection. This will do for N_3, but not for N_4, which would require one the be able to simultaneously accept and reject something.

If all else remains the same, this does not guarantee that the condition carries over to all sentences. (The induction proof breaks down in the case for the conditional). To ensure that is does, we have to change the falsity conditions for conditional to:

$$(A \supset B)\rho_w 0 \text{ iff for some } w' \text{ such that } wRw', A\rho_{w'}1 \text{ and } B\rho_{w'}0$$

Note that with these falsity conditions, the heredity condition no longer holds for all formulas, though it does hold for positive (negation-free) formulas.[10] Call the resulting logic, M.

As may be seen, it verifies the LEM. M is a paraconsistent logic, but not one that is acceptable to an antirealist since it verifies the LEM.

It should also be noted that the $\vee$-$\wedge$-$\neg$ fragments of N_4, M, and N_3 are the well known many-valued logics FDE, LP, and K_3, respectively.[11] (In these, the world structure becomes, in fact, irrelevant.) The first two of these are paraconsistent, but the second is ruled out for antirealist purposes because it verifies the LEM. FDE, however, is a perfectly acceptable paraconsistent antirealist logic.

10.5 Quantified Intuitionist Logic

So far, we have considered only propositional logics. Do the considerations carry over once we add quantifiers?

A Kripke interpretation for first order intuitionist logic is a structure $\langle W, R, D, \nu\rangle$. W and R are as in the propositional case. For every $w \in W$, D_w is a set of objects, subject to the constraint that:

$$\text{if } w_1 R w_2 \text{ then } D_{w_1} \subseteq D_{w_2}$$

The domain contains all those things we have constructed at that stage. We may construct new objects later, but what has been constructed stays constructed. For every constant, c, in the language, $\nu(c) \in D_w$ for all $w \in W$; and for every n-place predicate, P, and $w \in W$, $\nu_w(P) \subseteq D_w^n$, subject to the constraint that:

$$\text{if } w_1 R w_2 \text{ then } \nu_{w_1}(P) \subseteq \nu_{w_2}(P)$$

which is now the appropriate form of the heredity constraint.

Truth values (1, 0) at worlds are assigned to atomic formulas by the conditions:

$$\nu_w(Pa_1 \dots a_n) = 1 \text{ iff } \langle \nu(a_1), \dots, \nu(a_n)\rangle \in \nu_w(P)$$

[10] This means that the logic is not closed under uniform substitution. Closure can be regained by dropping the heredity condition for propositional parameters. This produces a system almost identical to that of [5, chap. 6]. The only difference is in the properties of the accessability relation.

[11] See [6, chaps. 7 and 8].

The truth conditions for the connectives are as in the propositional case. For the quantifiers:

$\nu_w(\exists x A) = 1$ iff for some $d \in D_w$, $\nu_w(A_x(k_d))$
$\nu_w(\forall x A) = 1$ iff for all w' such that wRw' and all $d \in D_{w'}$, $\nu_w(A_x(k_d))$

Here, we take the language to be augmented by a set of constants, k_d, such that k_d denotes d, and $A_x(c)$ is A with all free occurrences of x replaced by c. Note that the truth conditions for $\exists$ relate to just the instances at the world at issue, whilst those for $\forall$ relate to both it and all its future worlds.

As in the propositional case, one can show that:

if $w_1 R w_2$ and $\nu_{w_1}(A) = 1$ then $\nu_{w_2}(A) = 1$

Validity is defined, as in the propositional case, in terms of truth preservation at all worlds of all interpretations.[12]

These truth conditions naturally capture an appropriate antirealist understanding of the quantifiers. $\exists x A$ is verified at a stage just if some instance is. And $\forall x A$ is verified if every instance is verified whatever we go on to construct. Intuitively, this can happen only if we have a construction that applies to any object we come up with, d, to give a proof of $A_x(k_d)$. (If there is such a construction, then whatever object we construct at a later time, there will be a proof that it satisfies A. Conversely, if there is no such construction, then there is a possible development in which we find an object for which there is no proof.)

10.6 Quantified Logics of Constructible Negation

A first order version of N_4 is obtained from its propositional logic as for intuitionist logic. An interpretation is a structure $\langle W, R, D, \nu\rangle$. Interpretations are the same as for intuitionism, except that for every world, w, $\nu_w(P) = \langle E, A\rangle$, where $E, A \subseteq D_w^n$. (E and A are the extension and antiextension of P at w. I will write them as $\nu_w^+(P)$ and $\nu_w^-(P)$, respectively.) We now need a double heredity constraint:

if $w_1 R w_2$ then $\nu_{w_1}^+(P) \subseteq \nu_{w_2}^+(P)$
if $w_1 R w_2$ then $\nu_{w_1}^-(P) \subseteq \nu_{w_2}^-(P)$

And the relation ρ is defined in the natural way:

$Pa_1 \dots a_n \rho_w 1$ iff $\langle \nu(a_1), \dots, \nu(a_n)\rangle \in \nu_w^+(P)$
$Pa_1 \dots a_n \rho_w 0$ iff $\langle \nu(a_1), \dots, \nu(a_n)\rangle \in \nu_w^-(P)$

12 See [9, chap. 20].

The truth/falsity conditions for the connectives are as in the propositional case. For the quantifiers:

$\exists x A \rho_w 1$ iff for some $d \in D_w$, $A_x(k_d)\rho_w 1$
$\exists x A \rho_w 0$ iff for all w' such that wRw' and all $d \in D_{w'}$, $A_x(k_d)\rho_{w'} 0$
$\forall x A \rho_w 1$ iff for all w' such that wRw' and all $d \in D_{w'}$, $A_x(k_d)\rho_{w'} 1$
$\forall x A \rho_w 0$ iff for some $d \in D_w$, $A_x(k_d)\rho_w 0$

As in the propositional case, one can show that:

if $w_1 R w_2$ and $A\rho_{w_1} = 1$ then $A\rho_{w_2} = 1$
if $w_1 R w_2$ and $A\rho_{w_1} = 0$ then $A\rho_{w_2} = 0$

Validity is defined, as usual, in terms of truth preservation at all worlds of all interpretations.

Note that the falsity conditions for $\forall$ and $\exists$ are the reverse of what one might have expected. The quantifier $\exists$ relates to the world in question and all its future worlds; the quantifier $\forall$ relates just to the world in question. This is required to ensure that the heredity conditions hold for all formulas.

As for intuitionist logic, these truth/fasity conditions naturally capture an appropriate antirealist understanding of the quantifiers. $\exists x A$ is verified at a stage just if some instance is. It is falsified if every instance is falsified whatever we go on to construct. Intuitively, this can happen only if we have a construction that applies to any object we come up with, d, to give a proof of $\neg A_x(k_d)$. (If there is such a construction, then whatever object we construct at a later time, there will be a proof that it satisfies $\neg A$. Conversely, if there is no such construction, then there is a possible development in which we find an object for which there is no such proof.)

Dually, $\forall x A$ is falsified at a stage just if some instance is. It is verified if every instance is verified whatever we go on to construct. Intuitively, this can happen only if we have a construction that applies to any object we come up with, d, to give a proof of $A_x(k_d)$. (If there is such a construction, then whatever object we construct at a later time, there will be a proof that it satisfies A. Conversely, if there is no such construction, then there is a possible development in which we find an object for which there is no such proof.)

First-order versions of N_3 and M are obtained in exactly the same way. For N_3 we need the extra constraint that $\nu_w^+(P) \cap \nu_w^-(P) = \phi$. For M, the appropriate constraint is that $\nu_w^+(P) \cup \nu_w^-(P) = D_w^n$, and we modify the falsity conditions for the conditional as in the propositional case. It is not difficult to show that for N_3, no formula, A, is such that $A\rho_w 1$ and $A_{\rho_w} 0$; and for M, every formula, A, is such that either $A\rho_w 1$ or $A_{\rho_w} 0$.

Quantified N_3 is acceptable to an antirealist, but it is not paraconsistent. Quantified M is paraconsistent, but not acceptable to an antirealist because it validates the LEM. N_4 is both acceptable to an antirealist and paraconsistent.

Perhaps the most surprising thing about N_4 (and the other two logics) from the present perspective, is the following. Intuitionist logic verifies only the first three of the classical negation/quantifier exchange principles:

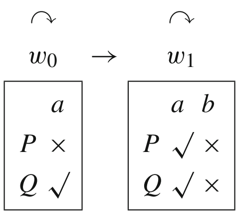
$$\forall x \neg Px \models \neg \exists x Px$$
$$\neg \exists x Px \models \forall x \neg Px$$
$$\exists x \neg Px \models \neg \forall x Px$$
$$\neg \forall x Px \models \exists x \neg Px$$

The fourth is invalid. The logics with relational semantics validate all four. Checking the first three is left as an exercise. Here is the fourth. Suppose that in world w of an interpretation $\neg \forall x Px \rho_w 1$. Then for some $d \in D_w$, $Pk_d \rho_w 0$. Hence, $\neg Pk_d \rho_w 1$, and $\exists x \neg Px \rho_w 1$.

It should be noted, though, that the intuitionistic invalidity:

$$\forall x(Pa \vee Qx) \not\models Pa \vee \forall x Qx$$

still fails, since this does not involve negation at all. (Here is a diagram of a standard counter-model:



The boxes give the extensions of P and Q at each world. The anti-extensions are irrelevant.)

10.7 Conclusion

We have now seen, as promised, that there are paraconsistent logics that are antirealism-friendly, and paraconsistent logics that are not. The examples examined were logics that deploy a relational semantics for negation. The main feature of the these logics for present purposes is that they treat truth and falsity even-handedly. This results in the validity of the Law of Double Negation and all the classical negation/quantifier exchange principles. These are a striking divergence from standard intuitionist logic, but perfectly defensible from an antirealist perspective, as we have seen.

Of course, there are many other paraconsistent logics, of widely different kinds.[13] To determine on which side of the realism/antirealism fence each sits requires its own investigation. Sometimes this will be obvious. For example, if the logic verifies the LEM, it is not going to sit on the antirealist side. Sometimes it will not be obvious. For example, do the ternary relation and the * function standardly employed in the semantics of relevant logics sustain an antirealist interpretation? This is a hard question, if for no other reason than that it is not clear what to make of these notions quite generally.[14] However, the present paper suffices to establish the general neutrality of paraconsistency on the realism/antirealism issue.

Acknowledgements Versions of this paper were given at the Universities of Melbourne, St Andrews, and Lille in the second half of 2007. I am grateful to the audiences in those places for their helpful comments.

References

1. Almukdad, A., and D. Nelson. 1984. "Constructible Falsity and Inexact Predicates." *Journal of Symbolic Logic* 49:231–33.
2. Mares, E. 2004. *Relevant Logic: A Philosophical Interpretation*. Cambridge: Cambridge University Press.
3. Nelson, D. 1949. "Constructible Falsity." *Journal of Symbolic Logic* 14:14–26.
4. Papineau, D. 1979. *Theory and Meaning*. Oxford: Clarendon Press.
5. Priest, G. 1987. *In Contradiction*. Dordrecht: Martinus Nijhoff; 2nd (extended) Edition, Oxford: Oxford University Press, 2006.
6. Priest, G. 2001. *Introduction to Non-Classical Logic*. Cambridge, MA: Cambridge University Press; revised as Part 1 of G. Priest, *Introduction to Non-Classical Logic: From If to Is* (Cambridge: Cambridge University Press, 2008).
7. Priest, G. 2002. "Paraconsistent Logic." In *Handbook of Philosophical Logic*, edited by D. Gabbay and F. Guenthner (2nd Edition), vol. 6, 287–393. Dordrecht: Kluwer.
8. Priest, G. 2006. *Doubt Truth to Be a Liar*. Oxford: Oxford University Press.
9. Priest, G. 2008. *Introduction to Non-Classical Logic: From If to Is*. Cambridge: Cambridge University Press.
10. Priest, G., and R. Routley. 1989. "The Philosophical Significance and Inevitability of Paraconsistency." In *Paraconsistent Logic: Essays on the Inconsistent*, edited by G. Priest, R. Routley, and J. Norman, ch. 18. Munich: Philosophia Verlag.
11. Read, S. 1994. *Thinking About Logic: An Introduction to the Philosophy of Logic*. Oxford: Oxford University Press.
12. Rumfitt, I. 2000. " 'Yes' and 'No.' " *Mind* 109:781–823.
13. Wansing, H. 2001. "Negation." In *The Blackwell Guide to Philosophical Logic*, edited by L. Goble, ch. 19. Oxford: Blackwell.

[13] For a survey of paraconsistent logics, see [7].

[14] The * semantics for negation are very closely related to the relational semantics, and in simple cases are interdefinable with them. See [6, §9.6], and [9, §22.5]. One might therefore reasonably expect the considerations concerning the relational semantics to carry over to the * semantics. For one interpretation of the ternary relation in terms of information, and so broadly sympathetic to an anti-realist reading, see [2, chap. 3].

Chapter 11
Type-Theoretical Dynamics

Exploring Belief Revision in a Constructive Framework

Giuseppe Primiero

11.1 Introduction

Various non classical models for dynamic reasoning have provided extremely fruitful results during the last two decades. The first aim of such frameworks is to interpret the basic standard properties of dynamic every-day reasoning: non-monotonicity, because conclusions earlier drawn on the basis of an insufficient set of informations can be rejected by new information obtained at a later stage; paraconsistency, because often the agent infers from contradictory informational contents; adaptive procedures, because reasoning is very often performed on the basis of an internal dynamics, which leads to the rejection of previously accepted consequences due to a better understanding or a modification of the starting set of premises. The standard reference for belief dynamics is to the AGM framework [1], and the large body of work that emerged from there.[1] In the present contribution, we explore the (constructive) type-theoretical model of knowledge as a framework to formalize non-monotonic procedures, by means of a proper interpretation of the logical notions at the basis of its formalism. This also provides a conceptual ground for a formal treatment of errors.

The type-theoretical dynamics provides a model defining independent items of information as beliefs, along with the usual epistemic definition of knowledge heritage of the constructive approach to logic. It provides the due operations on these states as the dynamic core of a knowledge process; it allows, moreover, an explanation of the procedures of information preference and it defines degrees of belief in terms of confidence and resistance to change. Our work proceeds as follows: we first introduce in Section 11.2 the type-theoretical framework in the context of conditions for developing an epistemic dynamics, together with the appropriate

[1] See [16, 33] for two of the most comprehensive treatments.

G. Primiero (✉)
FWO, Ghent University, Blandijnberg 2, Ghent, Belgium

IEG, Oxford University, Oxford, UK
e-mail: Giuseppe.Primiero@UGent.be

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_11, © Springer Science+Business Media B.V. 2012

epistemic explanations of the notions of belief, belief state and belief set, expectation; in Section 11.3 we interpret the revision operator, the Ramsey test and the Minimal Change Principle; in Section 11.4 an attempt is formulated to visualize the due operations for belief merging; we conclude in Section 11.5 with some further notions, useful in the context of extensions of the framework.

11.2 Conditions for Type-Theoretical Dynamics

The well-known idea at the basis of the classic AGM model[2] is to integrate new information within the starting knowledge frame of a rational agent, in order to consider which inconsistencies rise up, and how to deal with them. The basic way of treating with such a case is by accepting the *new information* and by changing the starting beliefs as little as possible (minimal change principle, MCP); the resulting belief state should be syntax independent. This model has been reformulated in different frameworks, each based on certain formal and conceptual constraints, and for each of them the meaning of "rational" is obviously depending on the type of logic used to represent the agent's procedures. Each of these models needs to express the conditions to formulate belief revision and non-monotonic procedures in terms of epistemic notions.

In the case of a type-theoretical frame, such conditions are easily presented. They can be informally introduced as follows:

1. to provide an appropriate formulation of the conceptual conditions (or equivalently of an order of priority) among the elements of the theory;
2. to describe an epistemic model in which an intutive interpretation of the basic notions of the dynamic of theories is provided, like e.g. for the notions of belief state and belief set, the revision operators and so on.

The first task must be obtained in a comparison with the usual descriptions of logical relations for contents of theories, namely by referring to one of the following methods:

- *epistemic entranchment*, originated with the AGM model, is an ordering among the sentences in a language in the form of a binary relation $\leq$, which tries to capture the importance of one of these sentences in face of a change[3];
- the *system of spheres* treats with sets of consistent complete theories and the order of relation among (parts of) them[4];
- a *preorder on models* as a structure equivalent both to entranchment and spheres.[5]

[2] Cf. [1].

[3] See [9].

[4] See [12].

[5] See [23].

These relations formulate the order of priority holding between propositional contents or (parts of) theories. This allows for any procedure of revision to be applied on ordered contents, i.e. depending on the relation of priority among the contents contained in the starting knowledge state of the agent.

In our type-theoretical framework, the definition of the conditional order is quite important and at the same time intuitive: the first problem is to define such relation among the judgemental (rather than propositional) contents of the theory, after which it appears natural to define the nature of doxastic states and their informational content in a completely new way; moreover, it is on the basis of such conditional order that revision operators and the epistemic notions involved (i.e. our second task) can be formulated. We proceed in doing this in the remaining of this section.

11.2.1 Type Theory as a Theory of Beliefs

The constructive type-theoretical approach to belief revision provides a quite developed framework to treat with beliefs, expectations and justifications.[6] Let us remember that the formal expressions of the theory are the standard categorical judgements[7]

$$a{:}A$$
$$a = b{:}A$$

and the dependent judgements:

$$a{:}A(x_1{:}A_1, \dots, x_n{:}A_n)$$
$$a = b{:}A(x_1{:}A_1, \dots, x_n{:}A_n).$$

Dependent judgements are formulated under assumptions contained in contexts (within brackets) and the primary condition for assumptions within contexts consists in the predicability of the types involved, in the form:

$$A{:}type.$$

[6] The framework presented in [4] is the first type-theoretical approach to belief revision. The authors stress the possibility to express explicitely an agent's beliefs as well as her justifications for these beliefs, by means of constructions which act as first-class citizens in the theory. The connection between types, their justifications, and the notion of belief is therefore based on the idea that any propositional content which in the type-theoretical formalization comes equipped with its instantiation (proof, instance) is intended as the content of a belief, which therefore results also justified. In the following, I will distinguish between contents as beliefs and contents as instances of knowledge.

[7] For all of these expressions, small letters represent instances or proofs, and capital letters represent types. The identity between instance and proof is based on the Curry-Howard isomorphism.

Our first aim is to explain this order of priority within judgements formation in relation to the models of theories for belief revision.

Notoriously, two models exist for belief revision theories: the first is represented by the "Foundations theories", in which one needs to keep track of justifications for one's beliefs, and beliefs are only accepted if justified; the second model is notoriously that of the "Coherence theories", in which beliefs are accepted on the basis of their coherence to other beliefs.[8] In the following, I shall maintain that the constructive type-theoretical approach to knowledge and belief falls among the Foundations Theories.

The first and major innovation provided by the type theory in the foundational perspective is obviously that it provides contents accompanied by appropriate justifications.[9] This has been considered one of the main elements in determining the nature of our (basic) beliefs,[10] and it has direct consequences on the definition of our belief states (see following subsection). If one understands type-theoretical judgements simply as judgements expressing beliefs, one will obtain a theory of "justified beliefs", because of the explicit formulation of constructions for types. On the other hand, a related problem, based also on the constructive foundation of the theory, is that it becomes extremely difficult to make sense of the nature of fallible theories: to accept the identity between judgemental forms and belief contents implies there are only *true beliefs*; no space is left for the formulation of *wrongly justified beliefs*, and *false beliefs* are interpreted as those propositional contents explicitely implying a contradiction: $a : A \rightarrow \perp$ (which are therefore rejected). This reading simply substitutes the usually understood notion of knowledge for the type theory by a notion of belief.[11] In the following, we choose to follow a different route: a proper notion of belief will be formally introduced as distinct from knowledge. Belief is considered to enjoy a weaker epistemic status than the strongly justified notion of knowledge. This explanation of belief contents has to be coherent with the epistemic definition of truth as presence of a proof, and it should express a content which can be submitted to revision and rejected when false. This will be obtained by interpreting the schema of conditions holding for the type-theoretical judgements, i.e. in terms of judgements not equipped with proper analytic constructions (proofs), or whose constructions are only assumed to be possessed.[12] This seems to express a more intuitive notion of belief as what (by definition) is not true, rather is *maintained to be true*. The interpretation of type-theoretical contexts as belief states identifies beliefs with contents which need to be processed before their acceptance, a method

[8] For this distinction, see especially [10, chap. 4, pp. 47–65].

[9] For a similar perspective in a different context, see [2].

[10] Cf. [14, pp. 22–23].

[11] This is the idea at the basis of the type-theoretical interpretation of [4].

[12] Let us consider e.g. the entry in the Oxford Dictionary (Paperback Edition, 2001):

> **belief**, n. **1** a feeling that something exists or is true, especially one without a proof. **2** a firmly held opinion. **3** (**belief in**) trust or confidence in. **4** religious faith.

that recalls the notion of *databases* in computer science and the idea of *acceptance system* from epistemology.[13]

The second issue that allows to treat our formalism among the Foundations Theories is the crucial role of priority relations in the building of contexts, which turns out to be another major issue in the description of belief generation and revision.[14] The formal structure introduces type theory as based on a calculus of contexts, and each content in such structure is strictly dependent on the formulation of the previous ones in the same context; moreover, a formal dynamics among contexts has already been developed,[15] and we are going to make extensive use of that too. The crucial step is to justify the interpretation of contexts as belief sets, which in turn let to interpret the relation of extension between contexts as the connection between different consecutive belief sets, providing a model for the standard operations of a belief revision theory, to be extended to the operation of merging. To accomplish this aim also means to import all the structural properties of contexts in the analysis of belief states, hence also gaining the prioritized structure and the possibility of defining revision operations on such structured contents. The next step is therefore the analysis of a belief state as a dependent judgement, with a knowledge declaration valid on the basis of an appropriate belief set.

11.2.2 Belief States and Belief Sets

The standard representation of doxastic states is sentential or propositional, i.e. beliefs are coded as formulas representing propositions. Following Martin-Löf's explanation of formulae in type theory, the propositional content A is embraced into the judgemental form *A is true*, obtained by abstraction on the formal expression $a:A$, saying that there exists a construction or proof a for A. This means that the formulation of the formal derivation (or construction) a justifies the truth of the concept or predicate expressed by the type A. This structure establishes therefore the truth of propositional contents on the basis of the proofs represented by closed derivations, and it defines in this way types as its formal objects.[16] Given this conceptual switch from simple propositions to the judgemental form predicating truth

[13] Cf. [20]. The definition of belief set for the type-theoretical framework provided above contains therefore as its core an internal definition of the informational contents of belief states at the agent's disposal. In [27] this is formulated in terms of the epistemic distinction between the "information" that a rational agent accepts in her system (and another agent could reject), and her "knowledge", as the stronger epistemic status of a certain content supported by a proof (amounting to Lehrer's distinction between "acceptance condition" and "knowledge").

[14] See [33, p. 135]:

> In accordance with the foundationalist philosophy of belief change, what gets revised in the first place is not *theories*, but rather *prioritized belief bases*.

[15] Cf. [31].

[16] See [21, 22].

of the propositional contents, closed derivations are the basic conditions for the latter, namely their *proof-conditions* (in the light of the Brouwer-Heyting-Kolmogorov interpretation of propositions). Judgements which provide proof-conditions for their propositional contents, are in the following intended as items of knowledge.

In addition to their strict proof conditions, it is also required that further conditions for judgements be formulated: those are expressed by the related *assertion conditions*, intended as the basis needed in order to formulate some knowledge content.[17] These assertion conditions for the type-theoretical formalism are represented by the expressions contained in contexts:

$$\Gamma = (x_1 : A_1, \ldots, x_n : A_n) : context.$$

A context collects assumptions and recall (implicitely) presuppositions for the knowledge content that is formulated. Contexts provide an informational content playing an essential role in the formulation of new judgements, and they let to formulate the required conceptual order between expressions of the theory.

The relation of dependency holding among the content of such contexts and the judgements derived from them, is based on the trasmission of knowability, rather than on truth-preservation as in classical models. The assumption on the knowability of certain propositional contents provides conditions for the knowability of further propositional contents. In this sense, it is the degree of knowledge and therefore the confidence in the truth involved by such contents that allows the interpretation of the notion of "belief" within the type-theoretical formalism. We shall interpret contextual contents to provide "independent items of information", not necessarily derived by actual processes of inference, and representing the foundation to build further contents.[18] In standard belief revision theories, these are called "default assumptions". The role of assumptions as independent items of information is simple to describe:

- an assumption of the truth of a certain propositional content is called an alethic assumption and it is of the form $(x{:}A)$;
- an assumption on the knowledge of a closed derivation making a certain content true is called an epistemic assumption and it is of the form $(a{:}A)$.[19]

Clearly, an alethic assumption can be considered an abstraction on an epistemic one, and viceversa, an epistemic assumption is the actual construction of an alethic one, thus being conceptually prior.[20] The formulation of different kinds of assumptions is essential for the epistemic counterparts of the notions of belief set, belief state and expectation. In the usual terminology of belief revision theories, the term belief set

[17] The conceptual and formal frame of conditions for type-theoretical judgements is fully presented and investigated in [25].

[18] See e.g. [33, chap. 1].

[19] See [25]. In both cases the presence of brackets refers to the use of the expressions in a context.

[20] For the notion of abstraction involved by proof processes and variables, see [28, 34].

(B_{set}) refers to the agent's set of beliefs closed under logical consequence, i.e. the set of actual beliefs and their consequences; the term belief state (B_{state}) refers instead only to the actual beliefs maintained by the agent in a certain knowledge process. To be more accurate: a belief base/state is a partial description of the world, which together with some inference operations generates a belief set. The beliefs in a base are therefore basic ones with an independent warrant, whereas those that follow from the base to obtain a state are the derived ones: the beliefs in the base are the foundation on which a belief set is built.[21]

In the type-theoretical model, contexts clearly represent belief states, that come additionally with appropriate justifications; the latter, in the alethic representation of assumptions, are abstracted. A B_{state} will therefore contain the agent's beliefs, and it will be part of a B_{set}, the latter providing a piece of knowledge with the appropriate conditions to express it:

$$B_{set} = \begin{array}{c} \overbrace{(\Gamma)}^{B_{state}} \\ \Downarrow \\ a{:}A \end{array}$$

This distinction is completely natural in the type theory, where a belief set can be based on different belief states (contexts), each providing a different collection of conditions on the knowledge expressed by the entire set. On the basis of the epistemic constructive interpretation here provided, the agent will be aware of further logical consequences only if equipped with the due constructions: this means that each B_{set} can be simply understood in terms of a knowledge state, containing a set of belief contents (assumptions and presuppositions) from which constructions are derived for new known propositional contents.

11.2.3 Beliefs and Expectations

On the basis of the above considered distinction between alethic and epistemic assumptions, the type-theoretical framework allows now to introduce a proper description for beliefs and expectations. These notions are usually explained[22] in terms of the agent's attitude towards the contents: a full belief is defined as a propositional content actually held true; when the same content is submitted to revision, one says that the expectation is contradicted, for example by new observations. This description can now be accounted in our epistemic schema by referring to different states produced by assumptions. Whereas the notion of full belief is clearly the proper equivalent of a justified belief, i.e. of a known content, the relation between

[21] Cf. [33, p. 22].

[22] See e.g. [11].

a belief and an expectation is completely based on the structure of dependent judgements.

A dependent judgement which is based on alethic assumptions $(x_1{:}A_1, \dots, x_n{:} A_n)$, has obviously a weaker status than a judgement depending on epistemic assumptions $(a_1{:}A_1, \dots, a_n{:}A_n)$, i.e. when the agent is actually able to show constructions that justify her assumptions. Hence, the belief set based on alethic assumptions expresses a conditional relation among the supposed truth of certain propositions and the truth of their consequence; when the belief set is based instead on epistemic assumptions, it expresses directly the knowability of contents. Assuming to possess a closed derivation for a certain propositional content implies therefore also that one assumes to be constrained by that derivation to a known content. This means that if the conditional relation really holds, the knowability of the conclusion requires the assumed knowability of the premises, and those known assumptions lead invariantly to that conclusion. Under these conditions, constructions should not be revised, up to the point an error is found in the construction itself, i.e. when a syntactical error is done. In other words, a dependent judgement holding under epistemic assumptions, really expresses an expectation (where by this word one also refers to something which is considered most likely to happen). A simple belief does not imposes the same epistemic constraints.

The difference between these forms of assumptions is therefore essential also in relation to the kind of errors involved: an alethic assumption $(x{:}A)$ can be rejected because either it is recognized to be meaningless, i.e. it fails to have a well-formed type-introduction $A : type$, or it assumes an old variable, i.e. the new assumption would correspond to an old construction. On the other hand, a revision performed on an epistemic assumption $(a : A)$ refers essentially to the syntactic structure of its construction a. A revision of a possible construction within the type A, recalls (at least implicitely) a transformation into the alethic correlative $(x : A)$. The correspondence between different forms of assumptions and the dynamic-theoretical distinction between beliefs and expectations has a second result, which is of the greatest importance: it allows to provide a correct type-theoretical interpretation of the Ramsey Test and therefore to explain the identity between revision and conditionals.[23]

11.3 Belief Revision

The wide range of non-classical approaches formulated for the problem of non monotonic reasoning, and in particular for the treatment of belief revision processes has produced a number of formal tools to treat with the cases of incoherence and revision, see e.g. [7]. Localization, representation and treatment of inconsistent data

[23] The first remarks on the interpretability of the Ramsey Test within type theory based on the correct understanding of the notion of assumption are due to Göran Sundholm.

are crucial properties that non explosive logics share. One of the first formalization by contraction on dispensable elements is introduced in [15], and a variation on the theme of inconsistent belief bases is given in [18]; in [36] the state of the beliefs on inconsistent information is considered a distinct element from the inconsistent data itself, other treatments admit inconsistencies by the use of a paraconsistent logic as in [32] or by an adaptive logic as in [29]; another model of retraction on inconsistent bases by default reasoning is given in [5].

Constructive interpretations cannot provide such a treatment for the constraint on consistent belief sets, but are richer than classical approaches in such that they work on a system of justifications. A constructive treatment has been restricted so far to some function of revision in classical structures, which acts constructively on propositional contents, see [23]. Our aim is to illustrate now an interpretation of the standard revision operators according to the conceptual frame of Constructive Type Theory introduced above.

The description of a revision procedure in the syntax of type theory is not so difficult to formulate,[24] but its development is based on a rather complex interpretation of the Ramsey Test and of conditionals. Notoriously, according to the Ramsey Test for conditionals,[25] a conterfactual conditional $A > B$ holds in a current body of knowledge K if and only if B is in K revised by A:

$$A > B \in K \leftrightarrow B \in K * A$$

with $*$ being one of the usual revision operators. From the point of view of the formal structure of type theory, a formulation of the test will require therefore an interpretation both of the conditional and of revision operations.

An interpretation of conditionals is included in the formulation of dependent judgements in terms of assertion conditions. Type theory has different ways to formalize the conditional relations for which the Ramsey Test expresses the identity with the revision operation. Among those, the case of a conditional statement (if A is true, then B is true) demands for its verification a dependent proof b of B provided that x is a proof of A or, in other words, the formulation of a construction for the type B under the alethic assumption that A *is true* $(b{:}B(x{:}A))$. This notion of conditional relation is the key term to interpret the Ramsey Test: if the identity proposed by the test is to hold also in the formal language of type theory, the notions of alethic and epistemic assumptions and their role in the interpretation of revision procedures has to be clarified.

---

[24] This formal frame has been spelled out entirely in [26]. The results therein contained will be here considered from a more conceptual point of view.

[25] See e.g. [8, 33].

11.3.1 Interpreting the Revision Operators

The operation of *revision*, $K{*}A$ in the AGM model consists notoriously in the introduction of a new belief maintaining consistency; in general, any belief change can be defined in that model by contraction and expansion.

In the type-theoretical model, the basic interpretation of the notion of belief in terms of contextual contents, leads to a concurrent description of expansion and contraction operations. Let us start by the first one. The primitive notion of expansion is defined in terms of a pair of operations:

- *expansion*, $K{+}A$: it corresponds to a modification of a belief state within a belief set, by interpreting a context into another one, an operation we denote as follows:

$$(K(\Gamma \leftarrow^{+} \Delta))$$

  where context Γ is expanded to context Δ in the belief set K by formulation of (at least one) new assumption $x_n : A_n$, with type A_n already contained in the starting context Γ;
- *update*, $K\diamond A$: it corresponds to a modification of a belief state within a belief set, by interpretating a context into another one, an operation we denote as follows:

$$(K(\Gamma \leftarrow^{\diamond} \Delta))$$

  where context Γ is updated by context Δ in the belief set K by formulation of (at least one) new type declaration $B : type$. On the basis of an update a related expansion by $(x{:}B)$ is allowed.

Expansion and update represent in our model the two main operations performed in order to enlarge a belief state. Each operation provides conditions for the derivation of new judgements within the belief set, i.e. for a possibly monotonic extension. This distinction is strictly related to the form and nature of contexts. Constructive Type Theory deals with judgement formation, and type declarations, i.e. judgements of the form $A : type$ are considered presuppositions for any judgement of the form $a{:}A$. In other words, a construction for type A, showing its constructive meaning, presupposes stating that A is a meaningful type apt for predication.[26] On the basis of this priority, the distinction between expansion and update is formulated: the former works on existing types formulating new alethic assumptions; the latter formulates new assumptions under new type declarations; hence, an update always implies implicitely or explicitely an expansion. The distinction here intended between expansion and update obviously reflects in a new way the standard distinction settled in [19], according to which the term *revision* refers to an epistemic change due to additional information produced in a static world, whereas an *update* is a change due to a variation happening in the world. These two operations in the

[26] The distinction between constructive meaning and meaningfulness is explained in more detail in [25, 27].

type-theoretical framework can also be linked to analytic and synthetic extensions of contexts:

- an expansion produces an analytic extension of the belief base, via the addition of one or more hypotheses falling within the given conceptual schema of the agent, i.e. within the types already provided, or by setting a definition for one such hypothesis;
- an update consists in a synthetic extension of the given conceptual frame, via introduction of one or more new predicable types, i.e. an extension of the meaning criteria of the body of knowledge.

The related derivation of new constructions within the belief set, is always to be intended as an analytic extension. A revision in terms of expansion or update can result nonetheless into an error; the explanation of this case requires the formulation of an appropriate type-theoretical operation:

- *contraction*, $K \underline{\circ} A$: is the result of removing one element in the belief set by means of a type-checking procedure performed backwards on the operations of

  1. derivation of new judgements;
  2. expansion of context;
  3. update of context.

By type-checking one identifies the failing belief by parsing any performed operations of revision; an error is categorized according to the corresponding operation.[27] Let us explore this issue further.

One starts the backward type-checking by analyzing performed derivations; if these are checked to be syntactically correct, the origin of a possible incoherent extension of a belief set is brought back to an operation of revision, i.e. by expansion or update. One first considers the formulation of (alethic/epistemic) hypotheses of the belief state in which an inconsistency is obtained including all of the involved type declarations: an error found at this step corresponds to an invalid formulation of an assumption within a type, hence a proof variable that cannot be verified. If any such construction is available, type-checking goes back to check udpates, i.e. the staged formulation of any new type (from the more to the less dependent one).[28]

---

[27] I am indebted to Göran Sundholm for much of my comprehension of the problem of error in this formal context, and for the consequences I am here ilustrating.

[28] A corresponding formal analysis has been introduced in [26, pp. 182–83], as the *Restricted Monotonicity Principle I/II*:

> **Restricted Monotonicity Principle I**: Let $(k_1 \cup \{i_1\})$ represent the *k-state* obtained by updating k_1 with the information expressed by i_1. It holds that $(k_1 \cup \{i_1\}) \preceq k_2$, i.e. that k_2 is coherently obtained from the expanded state $(k_1 \cup \{i_1\})$, iff $i_1 \subseteq k_1$.
>
> **Restricted Monotonicity Principle II**: Let $(k_1 \cup \{i_1\})$ represent the *k-state* obtained by updating k_1 with the information expressed by i_1. It holds that $(k_1 \cup i_1) \prec k_2$, i.e. k_2 is obtained from the updated state $(k_1 \cup i_1)$ and some of the truths derived could be incoherent with what known in (k_1), iff $i_1 \subseteq k_2$.

11.3.2 Ramsey Test and Minimal Change

Let us now come back to the interpretation of the Ramsey Test. Under the above displayed reading of the revision operations and provided the interpretation of conditional statements as judgements of the form $b{:}B(x{:}A)$, the formula $B \in K * A$ on the righthand side of the Ramsey biconditional, means that a certain formal judgement $b{:}B$ holds within a body of knowledge, whose context has been revised by an operation concerning the propositional content A. Such revision can be of the form $\leftarrow^{+} (x_n : A_n)$ – i.e. with A_n already contained in the starting context Γ and x_n a new assumption; or, rather. of the form $\leftarrow^{\diamond} (x_n{:}A_n(A_n{:}type))$ – i.e. an assumption formulated under a new type A_n. In both cases, the revision by contraction of A_n is defined by coming back to a primitive expansion: $\leftarrow^{+} (x_n{:}A_n \rightarrow \bot)$. This definition provides the type-theoretical interpretation of the Minimal Change Principle. The classical formulation of the principle is the following:

$$\neg A \notin K \rightarrow K * A = K + A \qquad \text{(MCP)}.$$

(MCP) says that, provided one does not hold the belief that A is not true, than the belief state is immediately revised by the belief that A is true. Some intuitive reasons challenge this principle,[29] but the kind of dynamics needed to solve these problems is difficult to simulate in the type-theoretical interpretation, especially in its constructive format. Our model provides in fact a far more restrictive meaning of contraction:

$$K \underline{\circ} A \leftrightarrow K(\Gamma) \leftarrow^{+} (x{:}A \rightarrow \bot) \qquad \text{(CONTR)}.$$

According to this definition of contraction, the rejection of a certain belief results only from an absurdity being implied by its assumption. On the other hand, not to possess such an implication from A to the absurdity does not imply the truth of the content at hand, at least until a proper construction for A is obtained. For this reason the standard formulation of the MCP does not hold in our model. This formulation of contraction is nevertheless equivalent to the principle of consistency for the AGM

---

where the informational state i_1 expresses an extension on the belief state, and the knowledge states (belief sets) k_1 and k_2 are the starting and the resulting belief set (before and after the extension). This reflects the suggestion formulated by [4], according to which type theory provides a more direct procedure to recognize "suspect beliefs": suspects are to be chosen among the elements of the context in which certain knowledge is derived, and the agent is allowed to choose which suspect to remove.

[29] If I do not hold the believe that a God does not exist, it does not mean necessarily that I hold the opposite believe that a God exists. For example, I could be agnostic, and therefore refuse to hold any of them; or I could be willing to maintain both of them, at different stages or even at the same time, by referring to different meanings of "existence" or "godness"; I could finally refuse to accept the predication of "existence" in connection to the subject God.

model, according to which the revision of K by A is incoherent only if it is proved in K that the negation of A holds:

$$(K * A = K \rightarrow \perp) \leftrightarrow \neg A \in K \qquad \text{(AGM-REV)}.$$

Under these conditions, the identity between conditional and revision proposed by the Ramsey Test is problematic because it results trivial, holding only for belief sets which are complete in respect to conditionals. Different solutions have been considered, like making the test weaker by adjoining preconditions, or by modifying the acceptance of conditionals and their relations to update. On the other hand, MCP is counterintuitive if interpreted for conditionals, because the introduction of new information can obviously change the antecedent of a conditional. The main task of the present section is to present the type-theoretical interpretation of the Ramsey Test, and to suggest how to make it non-trivial.

11.3.3 Type-Theoretical Ramsey Test

In its type-theoretical interpretation, the right-hand expression of the Ramsey identity means essentially that the revision $K * A$ provides conditions for a certain judgement B to be apt to be predicated. Given the translation of the conditional $A > B$ with the hypothetical judgement $b{:}B(x{:}A)$, the operation of revising the context in K by the content $x{:}A$ is therefore enough to be able to formulate a construction for B. This *type-theoretical Ramsey Test* should be formulated as follows:

$$b{:}B(x{:}A) \in K \leftrightarrow b{:}B \in K(\Gamma \leftarrow^{\diamond} (x{:}A)) \qquad \text{(ttRT)}.$$

The requirement that the revision operation be one of update ($\leftarrow^{\diamond}$) is meant to satisfy the following additional conditions in order $B \in K * A$ to be equivalent with the related conditional $A > B$:

1. the conditional $A > B$ is valid only under formulation of the due assertion-conditions for the antecedent, in particular its type-declaration $A{:}type$;
2. the formulation of the due assertion conditions for A does not directly imply that correct proof-conditions b for B have been formulated, in order the truth of B to be asserted.

Under these specifications, it is not always true that a certain conditional $b : B(x : A)$ holds in K *if and only if* B holds provided the minimal change of K with $(x : A)$. Out of the formalism, it makes completely sense to modify a certain body of knowledge with some content providing conditions for other knowledge, without in fact happening to know, or being able, or even willing to formulate the consequence of those conditions. The result of the biconditional is different (and it actually corresponds to the classical interpretation) if the conditional is intended as a proper consequence, i.e. an expression of the form "B is a consequence of A" or "A entails B", which in its complete epistemic formulation sounds "B is known if A is known". Such a formula says that extending a body of knowledge K by postulating an object $Proof(A)$, will imply $b{:}B$, i.e. an object $Proof(B)$. The way in which a corresponding revision on the body of knowledge K is performed, is completely different. One treats with extensions provided by means of an epistemic assumption, which happens to be the condition for another propositional content:

$$((A \rightarrow B)holds) \in K \leftrightarrow b{:}B \in K(\Gamma \leftarrow^{\diamond} (a{:}A)) \qquad \text{(ttRT2)}.$$

This modified version of the Ramsey Test (ttRT2) expresses the condition actually meant by the classic version: it accounts for the explicit knowledge of conditions allowing the formulation of a certain consequence. On the basis of the formulation of alethic assumptions, an agent being given with due conditions for a certain conditional, is not constrained to the explicit knowledge of their consequences; and to know the holding of a certain conditional does not mean to possess explicit knowledge of all the needed conditions.

11.3.4 A Model for Kripke Semantics

The result of the restricted monotonicity, which interprets extensions of (non-) monotonic reasoning under conditions of operations on contexts, can be compared with the connected notion of forcing when belief states are intended as nodes in a Kripke semantics. Notoriously, in such a semantics if a content is forced at a certain node, say i, it will be so at every further one j. In the type-theoretical structure this means, in the simplest case, that the result of a certain (non-)dependent derivation, i.e. one obtained under a (possibly empty) context Γ, is maintained at every next step whose extending context Δ is empty:

$$i \leq j, i \models a{:}A(\Gamma) \rightarrow j \models a{:}A(\Gamma \leftarrow (\Delta = \emptyset)).$$

This simple case needs to be modified in the case of an extension of the belief set based on a non-empty operation on context. The following equation considers the alternative case when Γ and Δ differ for at least one expression; in this case, the forcing of the given content at any later node strictly depends on the monotonicity of extensions of the previous context[30]:

$$i \leq j, i \models a{:}A(\Gamma) \rightarrow j \models a{:}A(\Gamma \leftarrow^{\diamond} \Delta) \leftrightarrow$$
$$(\Gamma \leftarrow^{\diamond} x{:}X = \Delta) \wedge (\neg(x{:}X \rightarrow \bot) \in \Gamma).$$

Therefore, monotonicity intended as

$$i \leq j, i \models a{:}A \rightarrow j \models a{:}A$$

[30] As in the previous cases, we consider the more general update, whose result is preserved by expansion.

is restricted under revision of contexts:

$$i \leq j, (i \models a{:}A(\Gamma) \rightarrow j \models a{:}A(\Delta)) \leftrightarrow i \models a{:}A(\Delta).$$

This analysis can be extended to interpret dynamic operations performed on contexts as the core property of type-theory[31]: the operations here considered allow to present the whole framework of type theory as a model of theory change, based on the constructive operation of preservation of knowability.

11.4 Belief Merging

The classical notion of merging in belief revision refers to the fusion of two or more sets of beliefs while maintaining consistency. Formal merging procedures typically refer to cases in which different sources provide the same information, inconsistent information, or in which their merging is intended to discover hidden or implicit beliefs (see e.g. [13]; for an overview see [6]).

As with belief revision, a constructive approach is more restricted and its analysis of merging processes advances with the same pro's and contra's mentioned in the previous sections. In particular, to interpret merging processes within the type-theoretical framework, one needs to provide a formalization which applies on contexts and extends the usual notion of monotonicity for expansions and derivations. The property of simple monotonicity in a (constructive) Kripke semantics applies to merging as follows: for every two stages $i \leq j$, there is an accessibility relation R_1, such that for every literal A such that $i \models A$, then $j \models A$; and if there are two stages i and j, and they are such that $i, j \leq k$, then there are two such relations R_1 and R_2 such that for every two literals A, B such that $i \models A, j \models A$, then $k \models A, B$:

$$\begin{tikzcd} i \arrow[dr, "R_1"'] & \\ & k \\ j \arrow[ur, "R_2"] & \end{tikzcd}$$

Within the type-theoretical approach, monotonicity has been interpreted on the derivations formulated on the basis of contexts. The stages i, j of the previous model of strict monotonicity can be now interpreted as distinct contexts Γ_i, Γ_j; A, B corresponds to (eventually distinct) constructions c_1, c_2 for

[31] See e.g. [24, 35].

corresponding propositional contents; and k is one and the same knowledge state obtained by those operations:

$$
\begin{array}{lll}
\Gamma_i & & \\
 & c_1 \searrow & \\
 & & k = \{a_1{:}A_1, \ldots, a_n{:}A_n\} \\
 & c_2 \nearrow & \\
\Gamma_j & &
\end{array}
$$

This formalizes a basic (without any protocol being defined) merging; moreover, this schema is always satisfied *if and only if* no operations of revision are performed on the contexts, and provided those contexts present both sufficient conditions for deriving the judgements in k. This is once again due to the conditions expressed by the restricted monotonicity, according to which only extensions executed without revisions on the informational contents of contexts lead with certainty to monotonic extensions of knowledge states. But a proper operation of merging between different belief states must be given under condition of revisions performed on the contexts of beliefs, i.e. precisely in the case when the Principle of Restricted Monotonicity may fail. Such a structure can be represented as a schema commuting under those explicit conditions of revision.[32] The two starting belief states Γ_i and Γ_j are now defined by operations of revisions r_1,r_2 (i.e. expansion or update) on a given state Γ[33]:

$$
\begin{array}{lll}
 & & \Gamma_i \\
 & r_1 \nearrow & \\
\Gamma & & \\
 & r_2 \searrow & \\
 & & \Gamma_j
\end{array}
$$

Provided these revisions, there is a merging if and only if a certain state k of the form $\{a_1 : A_1, \ldots, a_n : A_n\}$ is the result of (eventually distinct) derivations c_1, c_2 from Γ_1, Γ_j, i.e. that $\Gamma_i, \Gamma_j \models k$:

[32] It is easy to recognize the property illustrated in the following as the well-known "diamond property" for λ-calculi, which is usually defined as follows: a binary relation $\succ$ on the lambda terms satisfies the diamond property if for all terms M, M_1, M_2 for which $M \succ M_1$ and $M \succ M_2$ we have a term M_3 such that $M_1 \succ M_3$ and $M_2 \succ M_3$; the property is used especially for the reductions among lambda terms. See e.g. [3]. Martin-Löf in a talk given at the workshop "Mathematics, Algorithms and Proofs 2007", Lorentz Center, Leiden, made use of this basic property to show the type-theoretical interpretation of the inductive limit and logical operators in a comparison with Topos Theory.

[33] In the following schemas, we keep the direction of arrows to the right, but they actually correspond to revision operations among contexts, with left-oriented arrows, from the previous sections.

$$\begin{array}{ccccc} & & \Gamma_i & & \\ & r_1 \nearrow & & c_1 \searrow & \\ \Gamma & & & & k \\ & r_2 \searrow & & c_2 \nearrow & \\ & & \Gamma_j & & \end{array}$$

This means in turn that the operation of belief merging for the type-theoretical structure is interpreted in a rather strict way: for two belief sets to be mergeable, they need to provide sufficient assertion conditions to derive a unique knowledge state. This condition is too restrictive if it means that the merging is impossible provided the two contexts are equipped with contradictory elements. In fact, one interesting operation which often is presented in connection to merging is that of preference information. The basic case is the following: given two belief sets equipped with contradictory information among other, a preference operation should be performed, such that it makes possible to reject one of the two contradictory information (or both, if necessary), thus allowing the merging of the remaining informations in a new belief set.[34] A similar operation should be representable also in our type-theoretical frame and it has to be harmonized with the given formal representation of merging. To obtain this, one starts by considering a unique Γ: contradictory states are represented by expansions on such context leading respectively to $\Gamma_i = \Gamma \leftarrow^{+} (x{:}A)$ and to $\Gamma_j = \Gamma \leftarrow^{+} (x{:}A \rightarrow \bot)$. The merging of Γ_i, Γ_j is at this stage impossible. A revision is needed in order to reject one or both of the judgements: such operation can be performed following the already mentioned type-cheking procedure (any step requires that the checking has been performed on the previous ones, provided no error has occurred in the output):

1. analyze the set of inferences valid for k, performed on the basis of Γ_i:
   [1a.] if any judgement derivable from $k(\Gamma_i)$ is contradictory w.r.t. Γ, reject the extension provided by Γ_i;
   [1b.] if no contradiction is found, the checking procedure goes to the next step;
2. analyze the result of the substitution $(x/a : A)$ in Γ_i:
   [2a.] if an error is found in such a construction, reject the truth of type A, accept the extension of Γ_j;
   [2b.] if no error is found by assuming a construction in the type A, the extension provided by Γ_i is accepted and Γ_j will be submitted to contraction;
3. a counter-proof for the last step is represented by checking the presupposition $A{:}type$ in Γ_i: an error found at this stage means that the concept represented by that type is not admissible for predication, and indirectly implies the acceptance of Γ_j.

---

[34] Among other methods, a preference on contradictory belief bases that satisfies majority is defined in [30].

Once this type-checking is completed, one can accept one of the two extensions. The acceptance of the expansion given by Γ_i is expressed by the following schema:

$$\begin{tikzcd}
 & & \Gamma_i \arrow[dr, "c_1"] & \\
 & \Gamma_i \arrow[ur, "+(x:A)"] & & k \\
\Gamma \arrow[ur, "+(x:A)"] \arrow[dr, "+(x:A\to\bot)"'] & & \Gamma \arrow[ur, "+(x:A)"] & \\
 & \Gamma_j \arrow[ur, "\underline{\circ}(x:A\to\bot)"'] & &
\end{tikzcd}$$

Because Γ_i provides at least one new assumption, the schema is here completed by a new construction c_1 which is meant to lead to a new knowledge state k. The merging on the basis of Γ_j is instead given by the following schema, in which the expansion provided by the new assumption $(x : A)$ is rejected, and it is accepted the extension by means of $x : A \to \bot$:

$$\begin{tikzcd}
 & \Gamma_i \arrow[dr, "\underline{\circ}(x:A)"] & \\
\Gamma \arrow[ur, "+(x:A)"] \arrow[dr, "+(x:A\to\bot)"'] & & \Gamma_j \\
 & \Gamma_j \arrow[ur] &
\end{tikzcd}$$

The operation of merging allows therefore only a procedure for checking the admissibility of one side of two contradictory informations, and it does not allow a real internal dynamics of preference for informative contents at different stages.

11.5 Some Remarks

In this last section, the interpretation of type theory in the context of theories of change and the dynamics of knowledge system shall be considered in the light of some of its properties.[35] The conceptual order of conditions described at the beginning results here essential to interpret some of the most interesting features of standard models of dynamic reasoning.

35 These properties are partially extracted from the problems suggested in [17].

11.5.1 Admitting Beliefs

A first remark concerns the basic cognitive abilities of the rational agent implementing our system of belief change. Submitting beliefs to revision has been often represented in terms of finite belief bases (i.e. on calculi enjoying completeness and decidability) and finite ability of extensions. The type-theoretical model establishes stricter conditions on such abilities: this is obtained by the switch from a simple finite belief state, to a model in which beliefs are intended as conditions for proving knowledge. The criterion of predicativity for types, interpreted in terms of the definition of types from below by means of their constructions, represents the justification of the entire model of knoweldge. Beliefs, rather than being admitted with an upper limit bound, are restricted by a double criterion:

1. beliefs are admitted if meaning criteria are satisfied, by admission of predication aptness of the concepts involved;
2. beliefs are admitted if their formulation is needed by further judgements construction.

In this way, beliefs express a specific epistemic status for a rational agent who is carrying on a knowledge process, and extensions are accepted only if coherent with the starting knowledge base, i.e. only if the newly introduced expressions satisfy similar conditions. Moreover, the formulation of belief contents does not imply the logical closure of the related belief set, because the consequences require the formulation of due constructions. This analysis leads directly to a second topic, namely that of setting criteria for the degrees of beliefs.

11.5.2 Degrees of Belief

The degree of a certain belief, which is a natural property in every-day reasoning ("how much do you believe in this?") is formally expressed by the complementing notions of degree of confidence and resistance to change. The notion of degree of confidence is usually intended as a static criterion, held by the agent with respect to a certain belief: this notion is central in those approaches based on probabilistic accounts of beliefs. On the other hand, the notion of resistance to change is a dynamic criterion, directly depending on the degree given to a certain belief: this latter relation is mainly explained in the standard models by referring to the already introduced relation of epistemic entranchment. The first of the two relations holds in the models maintaining a distinction between sentences representing beliefs and those that do not (so called dichotomous models), whereas the distinction fails when resistance to change is called upon.

The epistemic structure described for the type theory can simply simulate the two relations here at hand, obviously in terms of the relation of provability for propositional contents and assertion-conditions for judgements. The main consequence is to set different epistemic states for categorical and dependent judgements. The static and the dynamic notions of degree of belief and resistance to change are therefore in this model completely connected, applying to the possible judgemental contents in the following way:

- a propositional content whose truth is asserted on the basis of its proof-conditions (a known content provided with its justification and related identity on proof terms, categorical judgement) provides the highest degree of confidence (certainty); its degree of resistance to change depends on the correcteness of the justification, therefore can vary only after revision of this construction;
- a propositional content whose truth is asserted depending on a context of assumptions (known content under believed justifications, dependent judgement) has a degree of resistance to change depending on the degree of confidence held by the agent with respect to the content of those assumptions;
- the contents of a certain context (belief contents, assumptions) provide a degree of confidence depending on the coherence and correcteness of the judgements derived on the basis of this context; their degree of resistance to change is the lowest, because these contents can always be submitted to revision, and it is also dependent on a certain degree of confidence in the meaningfulness the agent ascribes to the involved concepts (type-introductions).

The conceptual order establishes a fixed degree of confidence (comparable to value 1 in probabilistic models) for propositional contents whose proof-conditions are satisfied. This value is variable when referred to assertion-conditions of some proved judgement, i.e. when it is referred to the belief state of a certain knowledge process: in this case, the degree of confidence determines the resistance to change for dependent judgements and, viceversa, the degree of confidence in the derived judgements depends on the resulting coherence with the entire knowledge state and determines the resistance to change for their conditions. This latter value is the lowest, because assumptions are the most simple contents to revise, and they only depend on the meaningfulness of the types involved.

This model indirectly establishes also the degree of revision of contents, each content being epistemically determined by its proper justifications and conditions:

- a justified content is maintained as long as the justification is proved to hold, i.e. up to an error in the construction is found;
- a justified content given under conditions is maintained as long as those conditions are not proven false;
- conditions for knowledge are maintained as long as the derived content is coherent with the entire state, and the presuppositions of meaning for those conditions are satisfied.

The entire structure of type theory is essentially based on the conditional relation of justification, and the conceptual order provides in this way the conditions to revise contents. It seems clear that the constructive version of type theory, along with a number of other frameworks focusing on justifications,[36] provides a structure in which the notion of belief and a certain degree of dynamic for reasoning can be successfully simulated. The major property (and possibly its weakness) consists in a strong epistemic basis, which defines belief states in the context of a proper definition for knowledge, but which also imposes strict conditions on the possibility of revision, information preference and knowledge change.

Acknowledgements A first draft of this paper was written during a Research Fellowship at the University of Leiden (The Netherlands) in 2006. Thanks to Göran Sundholm and Maria van der Schaar for discussions related to various topics connected to this research. Later versions and this final form were drafted as a Postdoctoral Fellow at Ghent University (Belgium). Thanks to two referees for precious comments.

References

1. Alchourrón, C. E., P. Gärdenfors, and D. Makinson. 1985. "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions." *Journal of Symbolic Logic* 50:510–30.
2. Artemov, S., and E. Nogina. 2005. "Introducing Justification to Epistemic Logic." *Journal of Logic and Computation* 15(6):1059–73.
3. Barendregt, H. P. 1984. *The Lambda Calculus: Its Syntax and Semantics,* vol. 103 of *Studies in Logic and the Foundations of Mathematics* (Revised Edition). Amsterdam: North Holland.
4. Borghuis, T., F. Kamareddine, and R. Nederpelt. 2002. "Formalizing Belief Revision in Type Theory." *Logic Journal of the IGPL* 10(5):461–500.
5. Brewka, G. 1991. "Belief Revision in a Framework for Default Reasoning." In *The Logic of Theory Change Workshop, Konstanz,* vol. 465 of *Lecture Notes in Computer Science*. Berlin: Springer.
6. Cholvy, L., and T. Hunter. 1997. "Fusion in Logic: A Brief Overview." In *4th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty,* vol. 1244 of *Lectures Notes in Computer Science*, 86–95.
7. Gabbay, D., O. Rodriguez, and A. Russo. 2008. "Belief Revision in Non-Classical Logics." *The Review of Symbolic Logic* 1:267–304.
8. Gärdenfors, P. 1986. "Belief Revision and the Ramsey Test for Conditionals." *Philosophical Review* 95:81–93.
9. Gärdenfors, P. 1988. *Knowledge in Flux*. Cambridge, MA: Cambridge University Press.
10. Gärdenfors, P. 2005. *The Dynamics of Thought*. Synthese Library. Dordrecht: Springer.
11. Gärdenfors, P., and D. Makinson. 1994. "Nonmonotonic Inference Based on Expectations." *Artificial Intelligence* 65(2):197–245.
12. Grove, A. 1988. "Two Modelings for Theory Change." *Journal of Philosophical Logic* 17:157–70.
13. Halpern, J. Y., and Y. Moses. 1992. "A Guide to Completeness and Complexity for Modal Logics of Knowledge and Belief." *Artificial Intelligence* 52:319–79.
14. Hansson, S. O. 1994. "Kernel Contraction." *Journal of Symbolic Logic* 59:845–59.
15. Hansson, S. O. 1997. "Semi-Revision." *Journal of Applied Non-Classical Logics* 7(2): 151–75.
16. Hansson, S. O. 1999. *A Textbook of Belief Dynamics*. Applied Logic. Dordrecht: Kluwer.
17. Hansson, S. O. 2003. "Ten Philosophical Problems in Belief Revision." *Journal of Logic and Computation* 13:37–49.

[36] See e.g. [2].

18. Hansson, S. O., and R. Wassermann. 2002. "Local Change." *Studia Logica* 70:49–76.
19. Katsuno, H., and A. O. Mendelzon. 1991. "On the Difference Between Updating a Knowledge Base and Revising It." In *Principles of Knowledge Representation and Reasoning*, edited by J. A. Allen, R. Fikes, and E. Sandewall, 387–94. Los Altos, CA: Morgan Kaufmann.
20. Lehrer, K. 1990. *Theory of Knowledge*. London: Routledge.
21. Martin-Löf, P. 1987. "Truth of a Proposition, Evidence of a Judgement, Validity of a Proof." *Synthese* 73(3):407–20.
22. Martin-Löf, P. 1998. An Intuitionistic Theory of Types. In *Twenty-Five Years of Constructive Type Theory*, edited by G. Sambin and J. Smith, 127–72. Oxford: Oxford University Press.
23. Peppas, P., and M. A. Williams. 1995. "Constructive Modelings for Theory Change." *Notre Dame Journal of Formal Logic* 36(1):120–33.
24. Poole, D. 1988. "A Logical Framework for Default Reasoning." *Artificial Intelligence* 36: 27–47.
25. Primiero, G. 2004. "Presuppositions, Assumptions, Premises." Master's thesis, Universiteit Leiden.
26. Primiero, G. 2006. "Belief Revision in Constructive Type Theory." In *The Logica Yearbook 2005*, edited by O. Tomala and R. Honzík, 177–88. Filosofia: Czech Academy of Science.
27. Primiero, G. 2008. *Information & Knowledge*, vol. 10 of *Logic, Epistemology and the Unity of Sciences*. Berlin: Springer.
28. Primiero, G. 2009. "Proceeding in Abstraction. From Concepts to Type and the Recent Perspective on Information." *History and Philosophy of Logic* 30:257–82.
29. Primiero, G. 2010. *Prioritized Dynamic Retraction Function on Non-Monotonic Information Updates*, vol. 21 of *Studies in Logic*, 443–63. London: College Publications.
30. Primiero, G., and J. Meheus. 2008. "Majority Merging by Adaptive Counting." *Synthese* (KRA), 165(2):203–23.
31. Ranta, A. 1994. *Type-Theoretical Grammar,* vol. 1. Oxford: Clarendon Press, Oxford University Press.
32. Restall, G., and J. Slaney. 1995. "Realistic Belief Revision." Technical Report TRARP-2-95, Research School of Information Sciences and Engineering and Centre for Information Science Research, Australian National University, Canberra, Australia.
33. Rott, H. 2001. *Change, Choice and Inference—A Study of Belief Revision and Non-Monotonic Reasoning*. Oxford: Oxford Science.
34. Sambin, G., and S. Valentini. 1998. "Building Up a Toolbox for Martin-Löf Type Theory: Subset Theory. In *Twenty-Five Years of Constructive Type Theory*, edited by G. Sambin and S. Valentini, 221–44. Oxford: Oxford University Press.
35. Stalnaker, R. C. 1984. *Inquiry*. Cambridge, MA: MIT Press/Bradford Books.
36. Tamminga, A. M. 2001. "Belief Dynamics—(Epistemo)logical Investigations." PhD thesis, Universiteit van Amsterdam.

Chapter 12
Negation in the Logic of First Degree Entailment and *Tonk*

A Dialogical Study

Shahid Rahman

12.1 Introduction

Dialogical logic developed by Paul Lorenzen and Kuno Lorenz, was the result of a solution to some of the problems that arouse in Lorenzen' *Operative Logik* [15].[1] We can not discuss here thoroughly the passage from the operative to the dialogical approach, though as pointed out by Peter Schroeder-Heister, the insights of Operative logic had lasting consequences in the literature on proof-theory and still deserve attention nowadays.[2] Moreover, the notion of *harmony* formulated by the antirealists and particularly by Dag Prawitz has been influenced by Lorenzen's notions of *admissibility*; *eliminability* and *inversion*. However, on my view, the dialogical tradition is rather a rupture than a continuation of the operative project and it might be confusing to start by linking conceptually both projects together.

Dialogical Logic was suggested at the end of the 1950s by Paul Lorenzen and then worked out by Kuno Lorenz.[3] Inspired by Wittgenstein's *meaning as use* the basic idea of the dialogical approach to logic is that the meaning of the logical constants is given by the norms or rules for their use. This feature of its underlying

[1] Cf. Lorenz [14].

[2] Schröder-Heister [34].

[3] The main original papers are collected in Lorenzen and Lorenz [16]. A detailed account of recent developments can be found in Felscher [4], Keiff [9, 10], Keiff [11], Rahman [21], Rahman and Keiff [23], Rahman, Clerbout, and Keiff [2], Keiff [12], Fiutek, Rückert, and Rahman [5], Rahman and Tulenheimo [26], Rückert [31], Rückert [32]. For a textbook presentation (in English), see Redmond and Fontaine [29].

S. Rahman (✉)
Université de Lille, UdL3, CNRS, UMR 8163: STL, F-59653 Villeneuve d'Ascq, France
e-mail: shahid.rahman@univ-lille3.fr

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_12, © Springer Science+Business Media B.V. 2012

semantics quite often motivated the dialogical approach to be understood as a *pragmatist* semantics.[4]

The point is that those rules that fix meaning may be of more than one type, and that they determine the kind of reconstruction of an argumentative and/or linguistic practice that a certain kind of language games called dialogues provide. As mentioned above the dialogical approach to logic is not a logic but a semantic rule-based framework where different logics could developed, combined or compared. However, for the sake of simplicity and exemplification I will introduce only to the dialogical version of classical and intuitionist logics.

In a dialogue two parties argue about a thesis respecting certain fixed rules. The player that states the thesis is called Proponent (**P**), his rival, who puts into question the thesis is called Opponent (**O**). In its original form, dialogues were designed in such a way that each of the plays ends after a finite number of moves with one player winning, while the other loses. Actions or moves in a dialogue are often understood as *utterances*[5] or as speech-acts.[6] The point is that the rules of the dialogue do not operate on expressions or sentences isolated from the act of uttering them.[7] The rules are divided into particle rules or rules for logical constants (*Partikelregeln*) and structural rules (*Rahmenregeln*). The structural rules determine the general course of a dialogue game, whereas the particle rules regulate those moves (or utterances) that are requests (to the moves of a rival) and those moves that are answers (to the requests).

Crucial for the dialogical approach are the following points

1. The distinction between local (rules for logical constants) and global meaning (included in the structural rules)
2. The player independence of local meaning
3. The distinction between the play level (local winning or winning of a play) and the strategic level (global winning or existence of a winning strategy).
4. A notion of validity that amounts to winning strategy *independently of any model* instead of winning strategy for every model.
5. The notion of winning in a *formal play* instead of winning strategy in a model.

In order to highlight these specific features of the dialogical approach to meaning I will discuss the dialogical analysis of tonk, some tonk-like operators and the negation of the logic of first-degree entailment.

---

[4] Quite often it has said that dialogical logics has a *pragmatic* approach to meaning. I concede that the terminology might be misleading and induce one to think that the theory of meaning involved in dialogic is not semantics at all. Helge Rückert proposes the more appropriate formulation *pragmatistische Semantik* (*pragmatist semantics*).

[5] Cf. Rahman and Rückert [25, p. 111] and Rückert [31, chap. 1.2].

[6] Cf. Keiff [11].

[7] Tulenheimo [39].

12.2 Dialogical Logic and Meaning

12.2.1 Local Meaning

12.2.1.1 Particle Rules

In dialogical logic, the particle rules are said to state the *local semantics*: what is at stake is only the request and the answer corresponding to the utterance of a given logical constant, rather than the whole context where the logical constant is embedded.

- The standard terminology makes use of the terms *challenge* or *attack* and *defence*.[8] However let me point out that at the local level (the level of the particle rules) this terminology should be devoid of strategic underpinning.
- *Declarative utterances* involve the use of formulae, *interrogative utterances* do not involve the use of formulae

The following table displays the particle rules, where X and Y stand for any of the players **O** or **P**:

$\vee, \wedge, \rightarrow, \neg, \forall, \exists$	**Challenge**	**Defence**
X: $\alpha \vee \beta$	**Y**: ?-$\vee$	**X**: α *or* **X**: β (**X** chooses)
X: $\alpha \wedge \beta$	**Y**: ? $\wedge$ 1 or **Y**: ? $\wedge$ 2 (**Y** chooses)	**X**: α *respectively* **X**: β
X: $\alpha \rightarrow \beta$	**Y**: α (**Y** challenges by uttering α and requesting B)	**X**: B
X: $\neg\alpha$	**Y**: α	— (no defence available)
X: $\forall x\alpha$	**Y**: ?-$\forall x/k$ (**Y** chooses)	**X**: $\alpha[x/k]$
X: $\exists x\alpha$	**Y** ?$\exists$	**X**: $\alpha[x/k]$ (**X** chooses)

[8] See Keiff [11], Rahman, Clerbout, and Keiff [2]. Tero Tulenheimo pointed out that this might lead the reader to think that already at the local level there are strategic features and that this contravenes a crucial feature of the dialogical framework. Indeed, Laurent Keiff [11] introduced the terminology *requests* and *answers*. However the dialogical vocabulary has been established with the former choice and it would be perhaps confusing to change it once more.

In the diagram, $\alpha[x/k]$ stands for the result of substituting the constant k for every occurrence of the variable x in the formula A.

One interesting way to look at the local meaning is as rendering an abstract view (on the semantics of the logical constant) that distinguishes between the following types of actions:

(a) Choice of declarative utterances (=:disjunction and conjunction).
(b) Choice of interrogative utterances involving individual constants (=: quantifiers).
(c) Switch of the roles of defender and challenger (=: conditional and negation). As we will discuss later on we might draw a distinction between the switches involved in the local meaning of negation and the conditional.

Let us briefly mention two crucial issues to which we will come back later on

- **Player independence:** The particle rules are symmetric in the sense that they are player independent—that is why they are formulated with the help of variables for players. Compare with the rules of tableaux or sequent calculus that are asymmetric: one set of rules for the *true*(left)-side other set of rules for the *false*(right)-side. The symmetry of the particle rules provides, as we will see below, the means to get rid of tonk-like-operators.
- **Sub-formula property**: If the local meaning of a particle # occurring in φ involves declarative utterances, these utterances must be constituted by sub-formulae of φ.[9]

12.2.2 Global Meaning

12.2.2.1 Structural Rules

(SR 0) (starting rule)

The initial formula is uttered by **P** (if possible). It provides the topic of the argumentation. Moves are alternately uttered by **P** and **O**. Each move that follows the initial formula is either a request or an answer.

[9] This has been pointed out by Laurent Keiff and by Helge Rückert in several communications.

Comment: The proviso *if possible* relates to the utterance of atomic formulae. See formal rule *(SR 2)* below.

(SR 1) (No Delaying Tactics Rule)

Both **P** and **O** may only make moves that change the situation.

Comments: This rule should assure that plays are finite (though there might an infinite number of them). There are several formulations of it with different advantages and disadvantages. The original formulation of Lorenz made use of ranks; other devices introduced explicit restrictions on repetitions. Ranks seem to be more compatible with the general aim of the dialogical approach of distinguishing between the play level and the strategic level. Other non-repetition rules seem to presuppose the strategic level. One disadvantage of the use of ranks is that they make metalogical proofs quite complicated. Let us describe here the rule that implements the use of ranks.

- After the move that sets the thesis players **O** and **P** each choose a natural number n and m respectively (termed their repetition ranks). Thereafter the players move alternately, each move being a request or an answer.
- In the course of the dialogue, **O** (**P**) may attack or defend any single (token of an) utterance at most n (or m) times.

(SR 2) (formal rule)[10]

P may not utter an atomic formula unless **O** uttered it first. Atomic formulae can not be challenged.

The dialogical framework is flexible enough to define the so-called *material dialogues*, that assume that atomic formulae have a fixed truth-value:

(SR *2) (rule for material dialogues)

Only atomic formulae standing for true propositions may be uttered. Atomic formulae standing for false propositions can not be uttered.

(SR 3) (winning rule)

X wins iff it is Y's turn but he cannot move (either challenge or defend).

[10] See a discussion of this rule below in the commentaries about the dialogical notion of validity.

Global meaning

These rules determine the meaning of a formula where a particle occurs as a main operator in every possible play.

(SR 4i) (intuitionist rule)[11]

In any move, each player may challenge a (complex) formula uttered by his partner or he may defend himself against the last challenge that has not yet been defended.

or

(SR 4c) (classical rule)

In any move, each player may challenge a (complex) formula uttered by his partner or he may defend himself against any challenge (including those challenges that have already been defended once).

- Notice that the dialogical framework offers a fine-grained answer to the question: Are intuitionist and classical negation the same negations? Namely: The particle rules are the same but it is the global meaning that changes.

In the dialogical approach validity is defined via the notion of *winning strategy*, where winning strategy for X means that for any choice of moves by Y, X has at least one possible move at his disposal such that he (X) wins:

Validity (*definition*):

A formula is valid in a certain dialogical system iff **P** has a formal winning strategy for this formula.

Thus,

- α is classically valid if there is a winning strategy for **P** in the formal dialogue Dc(α).
- α is intuitionistically valid if there is a winning strategy for **P** in the formal dialogue Dint (α).

Examples: See Appendix 12.3.

[11] In the standard literature on dialogues, there is an asymmetric version of the intuitionist rule, called E-rule since Felscher [4]. For a discussion of this see appendices 1 and 2.

Comments on the dialogical notion of validity:

Helge Rückert [33] pointed out, and rightly so, that the formal rule triggers a novel notion of validity.[12] Validity, is not being understood as being true in every model, but as *having a winning strategy independently of any model* or more generally independently of any *material* grounding claim (such as truth or justification). The copy-cat strategy implicit in the formal rule is not copy cat of groundings but copy-cat of declarative utterances involving atomic formulae. Moreover one should add that there is the notion of *formal play*, that does not seem to correspond to nothing in model-theoretic approach: a formal play is not playing in a model.

In fact, one could see the formal rule as a process the first stage of which starts with what Laurent Keiff called *contentious* dialogues.[13] Contentious dialogues are dialogues where a player X utters one or more atomic formulae that are dependent upon a given ground and X is not prepared to put this ground into question—one can think of it as a claim of having some kind of ground (or a claim of truth) for it. Moreover, the antagonist is willing to concede this ground for the sake of the argument.[14] Now, if we would like to avoid to have the result that an atomic formula is true by the only reason that the player X stated it—that is, if we want to find a way out of contentious dialogues, then there are two possible ways:

- either we accept some principle of grounding external to the dialogue itself (and thus external to the interaction of the players)
  or
- we look for a player principle of grounding that is internal to the dialogue and dependent on the interaction of the players.

The first ways leads to material dialogues the second to formal ones

If we are willing to accept something like *material truth*, then we can think that the grounds upon which the atomic formulae depend are facts of the world and a grounded atomic formulae is a way to say that it is true. However, something more general might be thought too, such as true in virtue of some player independent ground. This is the basis on which the rule for material dialogues has been formulated.

Rückert pointed out that the formal rule establishes a kind of game where one of the players must play without knowing what the antagonist's justifications of the atomic formulae are. Thus, according to this view, the passage to formal dialogues relates to the switch to some kind of games with incomplete information. Now, if the ultimate grounds of a dialogical thesis are atomic formulae and if this is implemented by the use of a formal rule, then the dialogues are in this sense necessarily asymmetric. Indeed, if both contenders were restricted by the formal rule no atomic

---

[12] Talk at the worshop Proofs and Dialogues, Tübingen, Wilehm-Schickard Institut für Informatik, February 25–27, 2011.

[13] Cf. Clerbout, Keiff, and Rahman [2] and in Keiff and Rahman [13].

[14] Cf. Keiff and Rahman [13, pp. 156–57], where this is linked to some specific passages of Plato's *Gorgias* (472b–c).

formula can ever be uttered. Thus, we implement the formal rule by designing one player, called the *proponent*, whose utterances of atomic formulae are, at least, at the start of the dialogue restricted by this rule.

Apparently, the formal rule introduces an asymmetry in relation to the commitments of **O** and **P** particularly so in the case of the utterance of the conditional. Indeed, if **O** utters a conditional, then **P** is committed to utter a challenge that must at the end be based on atomic moves of **O**. If it is **O** that challenges a conditional no such commitment will be triggered. But it would be a mistake to draw from this fact the conclusion that the local meaning of the conditional is not symmetric. The very point of player independence is that it is a property of the meaning of the logical particles not of the dialogue as a whole where **P** is committed to the validity of the thesis. More precisely the asymmetry of the winning strategy is triggered by the semantic asymmetry of the formal rule. It is the possibility to isolate meaning (local and global) from validity commitments that allows dialogicians to speak of the symmetry of the logical constants and this prevents tonk-like operators from being introduced in the dialogical framework (see Section 12.2.3 below).

12.2.3 Play Level, Strategic Level and Tonk-Like-Operators

12.2.3.1 Strategic Level and Tableaux

As mentioned above in the dialogical approach validity is defined via the notion of *winning strategy*. A systematic description of the winning strategies available for **P** in the context of the possible choices of **O** can be obtained from the following considerations[15]:

If **P** is to win against any choice of **O**, we will have to consider two main different situations, namely

- the dialogical situations in which **O** has uttered a complex formula, and
- those in which **P** has uttered a complex formula.

We call these main situations the **O**-cases and the **P**-cases, respectively. In both of these situations another distinction has to be examined:

(i) **P** wins by *choosing* between two possible challenges in the **O**-cases or between two possible defences in the **P**-cases, iff he can win with *at least one* of his choices.
(ii) When **O** can *choose* between two possible defences in the **O**-cases or between two possible challenges in the **P**-cases, **P** wins iff he can win *irrespective* of **O's** choices.

[15] Clerbout [1] developed an algorithm that establishes the exact correspondence between tableaux and dialogical winning strategies. The algorithm is the most thorough result of the existing literature.

The description of the available strategies will yield a version of the semantic tableaux of Beth that become popular after the landmark work on semantic-trees by Smullyan [35], where **O** stands for **T** (left-side) and **P** for **F** (right-side) and where situations of type **ii** (and not of type **i**) will lead to a branching-rule.

(**P**)-***Chooses***	(**O**)-***Chooses***
$(\mathbf{P})\alpha \vee \beta$	$(\mathbf{P})\ \alpha \wedge \beta$
- - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - - - - - -
$<\mathbf{O}?>\ (\mathbf{P})\alpha$ $<\mathbf{O}?>\ (\mathbf{P})\beta$	$<\mathbf{O}?\wedge 1>\ (\mathbf{P})\alpha \mid\ <\mathbf{P}?\wedge 2>\ (\mathbf{P})\beta$
The expressions of the form <X...> constitute interrogative utterances	*The expressions of the form <X...> constitute interrogative utterances*
$(\mathbf{O})\alpha \wedge \beta$	$(\mathbf{O})\ \alpha \vee \beta$
- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - - - - - -
$<\mathbf{P}?\wedge 1>\ (\mathbf{O})\alpha$ $<\mathbf{P}?\wedge 2>\ (\mathbf{O})\beta$	$<\mathbf{P}?>\ (\mathbf{O})\alpha \mid\ <\mathbf{P}?>\ (\mathbf{O})\beta$
$(\mathbf{P})\alpha \rightarrow \beta$	$(\mathbf{O})\ \alpha \rightarrow \beta$
- - - - - - - - - - - - - -	- - - - - - - - - - - - - - - - - - -
$(\mathbf{P})\alpha$ $(\mathbf{O})\beta$	$(\mathbf{P})\ \alpha$ $(\mathbf{O})\ ?\ldots \mid (\mathbf{O})\beta$ *(Opponent has the choice between counterattacking or defending)*
No choice	***No choice***
$(\mathbf{P})\ \neg\alpha$	$(\mathbf{O})\ \neg\alpha$
- - - - - - - - - - - -	- - - - - - - - - - - - -
$(\mathbf{O})\alpha$	$(\mathbf{P})\ \alpha$

However, tableaux are not dialogues. The main point is that dialogues are built up bottom up, from local semantics to global semantics and from global semantics to validity. This establishes the priority of the play level over the winning-strategy-level. The levels are to be thought as defining an order. From the dialogical point of view, to set the meaning of the logical constants via validity is like trying to define the (meaning) moves of the king in the game of chess by the strategic rules of how to win a play. Neither semantic tableaux nor sequent calculus give priority to the play level. The point is not really that sequent calculus or tableaux do not have a play level, if with this we mean that one could not find the steps leading to the proof though there is one. What distinguishes the dialogical approach from other approaches is that in the other approaches—if there is something like a play level—the play level is ignored: the logical constants are defined via the rules that define validity.[16] The dialogical approach takes the play level as the level where meaning

[16] The point that other systems have also a play level has been stressed by Luca Tranchini in the workshop Workshop Amsterdam/Lille: *Dialogues and Games: Historical Roots and Contemporary Models*, February 8–9, 2010, Lille.

is set and on the basis of which validity rules should result. Within the dialogical approach, the more basic step of meaning at the play level is the setting of player-independent particle-rules (i.e. symmetric rules): the difference between **O** (**T**)-rules and the **P** (**F**)-rules results from the asymmetry introduced by the formal rule at the strategic level. These considerations lead us to tonk. One can build tableaux-rules for tonk and tonk-like operators but, from the dialogical point of view, they have no semantic underpinning.

12.2.3.2 Tonk, Tunk and Black Bullet

Let discuss our point by analysing first the case of tableaux-rules for Prior's original tonk. That is, let us assume that we start not by laying down the local meaning of tonk; but by specifying how a winning strategy for tonk would look like with the help of **T**(left)-side and **F**(right)-side tableaux-rules (or sequent-calculus) for logical constants.

Prior's tonk takes half of the rule that delivers the grounds for the assertion of a disjunction (half of the introduction rule) and half of the inference rule for the conjunction (half of the elimination rule). This renders the following tableaux version for the undesirable tonk[17]:

(**O**)[(**T**)] αtonkβ	(**P**)[or (**F**)] αtonkβ
- - - - - - - - - - - - - -	- - - - - - - - - - - - - - - -
(**O**)[(**T**)]β	(**P**)[(**F**)]α

From the dialogical point of view, the rejection of tonk is linked to the symmetry condition of the particle rules that cannot be fulfilled for tonk. Indeed; the defence must yield a different formula, namely the tail of tonk if the defender is **O** and the head of tonk if the defender is **P**:

O-tonk	**Challenge**	**Defence**
O: αtonkβ	**P**: ?	**O**: β

P-tonk	**Challenge**	**Defence**
P: αtonkβ	**O**: ?	**P**: α

This means, that the attempted particle-rule for tonk is player-dependent, and this should not be the case. The point is that in dialogues tonk-like operators are rejected because there is no symmetric particle rule that justifies the tableaux-rules designed for these operators.

[17] If we apply the cut-rule it is possible to obtain a closed tableau for **T**α, **F**β for any α and β.

Let us discuss the further example of the following tableaux-rules for a tonk-like operator that we call *tunk*:

(**O**)[(**T**)] αtunkβ	(**P**) [or (**F**)] αtunkβ
- - - - - - - - - - - - - -	- - - - - - - - - - - - - - - -
(**O**)[(**T**)] α	(**P**)[(**F**)] α
(**O**)[(**T**)] β	(**P**)[(**F**)] β

Such a constant, when added to the standard tableaux-rules of, say, classical logic renders proofs for

αtunk¬α, and for
¬(αtunk¬α)

Let us attempt to define a player independent particle rule for tunk. Let us thus assume that for a given player **X** that uttered *A*tunk*B* the challenge (if it should somehow meet the tableaux-rules) must be one of the following:

(1) (**Y**) show me the left side, and (**Y**) show me the right side. Here it is the challenger who has the choice;
(2) (**Y**) show me at least one of the both sides. Here it is the defender who has the choice.

Now whatever the options are, one of them will clash with one of the tableaux-rules described above:

- If we take option one, the challenger **O** has the choice and this should yield at the *strategic level* a branching on the **P**-rule and no branching on the **O**-Rule.
- If we take option 2, the defender **O** has the choice and this should produce a branching at the *strategic level* on the **O**-rule and no branching on the **P**-Rule.

Here are the correspondent (*un-dialogical*) asymmetric particle rules:

O-tunk	**Challenge**	**Defence**
O: αtunkβ	**P**: ? 1 or **P**: ? 2 (**P** chooses)	**O**: α respectively **O**: β

P-tunk	**Challenge**	**Defence**
P: αtunkβ	**O**: ?	**P**: α or **P**: β (**P** chooses)

The asymmetry comes here from the fact that **P** chooses in both cases. This is not a surprise since tunk has been designed extending the rules of tonk and dialogically speaking tonk takes those part of the rules where the Proponent have the choice.

The following formulation stresses this point

X-tunk	**Challenge**	**Defence**
X: αtunkβ	**Y**: ?	**X**: α respectively **X**: β (**P** chooses)

The point is that the tableaux-rules for tonk and tunk are not based on particle rules that are player-independent and are thus not apt to render local meaning. Moreover the only rules tableaux have are player dependent. That is why, according to the dialogical analysis, external criteria such as harmony; have to be introduced in order to reject tonk-like operators. The tableaux-tunk rules allow proving formulae that correspond to no winning strategy of **P**.

Stephen Read introduced a different kind of pathological logical constants—called *black-bullet* and the dual *white-bullet*.[18]

Let us have the following tableaux-rules for the logical constant *black-bullet*, that can be thought as a kind of a cero-adic operator and that says of itself that it is false:

$$\frac{(\mathbf{F})\bullet}{\begin{array}{c}(\mathbf{T})\bullet\\(\mathbf{F})\bot\end{array}}$$

$$\frac{(\mathbf{T})\bullet}{(\mathbf{F})\bullet \mid (\mathrm{T})\bot}$$

Blackbullet is certainly pathological:

18 Read [27, 28]

The tableaux-rules for black-bullet described above deliver closed tableaux for both ¬• and •.

From the dialogical point of view we can formulate symmetric particle rules for •:

•	Challenge	Defence
X: •	**Y**: ?	**X**: ¬•

Or

•	Challenge	Defence
X: •	**Y**: •	**X**: ⊥

Furthermore, the dialogical analysis of this particle allows two approaches:

(i) If we put the emphasis in the fact that • is an operator, then a dialogue with this operator as a thesis will generate an infinite game,
(ii) if we stick to a semantics that is complete in relation to the tableaux-rules then • has a double nature, namely, on one side it is an operator that can be challenged and on the other side it is an atomic formula and as such should follow the formal rule. This double nature could be rendered by adding a special structural rule like the following:

Black-bullet formal rule: ***O*** *can challenge* • *iff he has not uttered it before.*

The particle rules for black-bullet make it apparent that • is part of the challenge and defence moves and thus contravenes the sub-formulae property mentioned above.

Is there then any limit to the dialogical framework for the introduction of logical constants? As well known, since the work of Dag Prawitz,[19] the natural deduction framework provides some criteria for the introduction of logical constants, which, as mentioned above, are rooted in Lorenzen's inversion principle and are known as *harmony*. In the natural deduction framework there are only two sets of rules and thus might be thought as setting the meaning and the other as setting inferences that vehicle this meaning. In such a framework *local soundness* or reducibility says that any derivation containing, say, an introduction of a logical constant followed immediately by its elimination can be turned into an equivalent derivation without this detour. It is a check on the *strength* of elimination rules: they must not be so strong that they include knowledge not already contained in its premises. Dually,

[19] Prawitz [17, chap. IV]. See too Sundholm [36–38], Read [27, 28].

local completeness says that the elimination rules are strong enough to decompose a connective into the forms suitable for its introduction rule. It is still an open question if harmony should or not be based on the introduction rules as setting meaning rather than in the elimination rules.

The nice point is that in the dialogical framework we have in fact several different sets of rules. I will separate two of them, those that set the meaning (particle + structural rules) and those responsible for the inferences setting validity, that is the winning-strategies described by an adequate tableaux or sequent calculus. Thus, the version of harmony appropriate to dialogical logic is the local soundness and completeness of the calculus purported to describe the winning strategies of a given system. In the particular cases of tonk, tunk and bullet mentioned above the point is that the tableaux-rules are unsound in relation to the semantics established at the play level by the joint collaboration of the particle and structural rules. The basis of the latter set of rules is the player-independence and the sub-formulae property of the particle-rules. The pathological operators are rejected since the tableaux-rules with the help of which they have been described are too strong, the tableaux-rules prove more than the dialogical semantics allow. What we must do then in order to test if an operator is or not a tonk-like operator is to prove soundness and completeness in relation to the dialogical system described above. According to this argument, such metalogical proofs are crucial for the general means of the dialogical framework as a framework. In fact, Lorenzen and Lorenz started this path, which must now be worked out for the recent dialogical systems recently developed.

But can we not produce tonk-like operators only at the play-level? For example, let us take a particle rule for a tonk-like operator that gives as an answer to the challenger, say the head of tonk. Such a rule is possible, but the result is not very harmful, as it amounts to the introduction of an operator that is equivalent to any formula. Thus, it is fully redundant. A similar variation of tunk seems hard to find. Indeed, if we fix the meaning of tunk, say, by establishing that the defender has the choice then the particle rule will be exactly the one for disjunction. Another open question is about the limits on the structural rules. Can we freely combine a structural rule with the introduction of an arbitrary particle? The results coming from linear logic and substructural logics seem to indicate that there are many delicate interrelations to be taken into consideration. Deeper research is still due, however let us fix some points towards *dialogical harmony*:

Dialogical Harmony

1. Particle rules must be player-independent.
   This should also be understood that the particle rules should be defined independently of who is the player that is restricted by the formal rule.
2. Particle rules must fulfil the sub-formula property.
3. (*The particle rule of a logical constant must be given independently of the inner structure of the formula in which this logical constant occurs as a main operator.*)
4. Global meaning must be player-independent.

5. This assumes that within the structural rules a level of global meaning can be distinguished. This also assumes that the global meaning does not "undo" the player-independence of the particle rules.
6. Appropriate tableaux systems must be build up bottom up.
7. In other words; those tableaux systems (or sequent calculi), that render a proof theory for a given dialogical semantics must be sound and complete in relation to the latter.

The third condition can be contested as being too strong and is crucial for the discussion of the so-called "dual negation". In fact, a contravention to the third condition, as will see below does not seem to trigger tonk-like operators.

Can we establish a kind of dialogical Harmony theorem?

- The particle # is trivializing iff there are no symmetric particle rules for # (with sub-formula property).

Well, what we can do for the moment is to prove the following:

Partial—Dialogical-Harmony-Lemma I (PDL-1):

(PDL-1.1) If there is a trivializing particle # such that the tableaux-rules—constituted by two lines—(with sub-formula property)—have the following form:

(**T**) α[#]	(**F**) α[#]
- - - - - - - - - - - - -	- - - - - - - - - - - - - - -
(**T**) β	(**F**) β
(**T**) γ	(**F**) γ

Then there are no symmetric particle rules (with sub-formula property).

Proof: By contraposition, if there are symmetric particle rules for #, then the tableaux resulting from the winning strategies based on that particle rules do not correspond to the form described above.

Let us start with the case where the tableaux are constituted by two lines:

If there were symmetric particle rules for #, then defences and challenges must be player independent.

Let us thus assume that for a given player **X** that uttered α[#] the challenge (if it should somehow meet the tableaux-rules) must be one of the following:

(1) (**Y**) show me the left side, and (**Y**) show me the right side. Here it is the challenger who has the choice;
(2) (**Y**) show me at least one of the both sides. Here it is the defender who has the choice.

Now whatever the options are, one of them will clash with one of the tableaux-rules described above when we replace the variables by players:

- If we take option one, the challenger **O** has the choice and this should yield at the *strategic level* a branching on the **P**-rule and no branching on the **O**-Rule.
  That is, we should have

(**P**) α[#]

- - - - - - - - - - - - - - - - - - -

(**P**) β| (**P**) γ (where the challenger **O** has the choice)

(**O**) α[#]

- - - - - - - - - - - - - - - - - - -

(**P**) β

(**P**) γ (where the challenger **P** has the choice)

- If we take option 2, the defender **O** has the choice and this should produce a branching at the *strategic level* on the **O**-rule and no branching on the **P**-Rule.

That is, we should have

(**O**) α[#]

- - - - - - - - - - - - - - - - - - -

(**O**) β| (**O**) γ (where the defender **O** has the choice)

(**P**) α[#]

- - - - - - - - - - - - - - - - - - -

(**P**) β

(**P**) γ (where the defender **P** has the choice)

The case for one-line-tableaux is simpler:

(PDL-1.2) If there is a trivializing particle # such that the tableaux-rules—constituted by one line—have the following form:

(**T**) α[#]	(**F**) α[#]
- - - - - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - - - - - -
(**T**) β	(**F**) γ

(where β is different from γ)

Then there are no symmetric particle rules (with sub-formula property).

If there were symmetric particle rules for # then the defence must be constituted by one sole sub-formula that is uttered player-independently and the correspondent tableaux must be then the following

(**O**) α[#]

- - - - - - - - - - - - - - - - - - -

(**O**) β

(**P**) α[#]

- - - - - - - - - - - - - - - - - - -

(**P**) β

The corresponding lemma from right to left is:

If there are no symmetric particle rules for # (with subformula property) then there are trivializing tableaux-rules for this very particle.

The proof is still lacking

Black-bullet-like operators require a further lemma:

Partial—Dialogical-Harmony-Lemma 2: If there is a trivializing particle # such that the tableaux-rules have no sub-formula property then there are no particle rules with subformula property.

Proof: Left to the reader.

12.3 The Dialogical Meaning of Negation and the Logic of First-Degree Entailment

12.3.1 The Logic of First-Degree Entailment

The main of this section of the paper is to discuss the meaning of negation involved in the logic of First Degree Entailment (FDE) in the dialogical framework.[20] We will first present a proof theoretical and a relational semantics for FDE-negation.

12.3.2 Hintikka's Trees for Enquiry Games and FDE-Negation

By Jaakko Hintikka, Ilpo Halonen and Arto Mutanen [8] initiated a series of papers on Interrogative Logic and developed a framework called Enquiry Games.[21] The main point of Enquiry Games is the combination of interrogative moves with deductive moves. A crucial point of the deductive moves is that they are determined by the use of a tableaux-system that Hintikka and collaborators call *a variant of Beth's tableau for classical logic*.[22] This is not exactly so. If the two extra non-standard

[20] Laurent Keiff, in his PhD [11], developed a similar dialogical analysis of negation within a speech-act framework without linking it to FDE-negation. In fact Keiff studies different negations and he seems to suggest that what I identify as being a FDE negation is not really an operator.

[21] Hintikka, Halonen, and Mutanen [8, pp. 48–50].

[22] Hintikka, Halonen, and Mutanen [8, p. 48].

closure rules of Hintikka trees are deleted, the result is a system that is sound and complete in relation to Anderson and Belnap's logic of first degree entailment (FDE). Thus, one should say that Hintikka's trees for Enquiry Games are rather a variation of trees for FDE.

The semantic tree system associated to Enquiry Games is the following[23]

(T)-*Cases*	**(F)-*Cases***
(**T**)$\alpha \vee \beta$ - - - - - - - - - - - - - - - - - - - (**T**)α| (**T**)β	(**F**)$\alpha \vee \beta$ - - - - - - - - - - - - - - - - - (**F**)α (**F**)β
(**T**)$\alpha \wedge \beta$ - - - - - - - - - - - - - - (**T**)α (**T**)β	(**F**) $\alpha \wedge \beta$ - - - - - - - - - - - - - - - - - - - (**F**)α| (**F**)β
(**T**) $\sim(\alpha \wedge \beta)$ - - - - - - - - - - - - - - (**T**)$\sim\alpha$| (**T**) $\sim\beta$	(**F**) $\sim(\alpha \wedge \beta)$ - - - - - - - - - - - - - - - - (**F**) $\sim\alpha$ (**F**) $\sim\beta$
(**T**) $\sim(\alpha \vee \beta)$ - - - - - - - - - - - - - - (**T**) $\sim\alpha$ (**T**) $\sim\beta$	(**F**) $\sim(\alpha \vee \beta)$ - - - - - - - - - - - - - - - (**F**) $\sim\alpha$| (**F**) $\sim\beta$
(**T**) $\sim\sim\alpha$ - - - - - - - - - - - - - (**T**)α	(**F**) $\sim\sim\alpha$ - - - - - - - - - - - (**F**)α

Notice that negation does not produce "switch of sides"—the subformula(e) of a **F**(**T**)-formula will always be **F**(**T**)-signed. Furthermore, there is no rule for the conditional, that is, the most typical crossing-sides connective.

Let us talk about the closure rules. If we only have the standard rules:

- A branch for a Hintikka-tree is closed if it contains atomic formulae of the form (**T**)p and (**F**)p (for atomic p).
- A Hintikka-tree is closed if all its branches are closed.

[23] Actually Hintikka, Halonen, and Mutanen [8, pp. 48–50] present a sequent-calculus version. Priest [18, pp. 141–44] displays essentially the same tableau as the those of Hintikka, Halonen, and Mutanen, apparently without knowing them. However Priest calls them explicitly *FDE-tableaux*.

The resulting logic is not at all classical logic. In fact, it should be easy to see that we can not close the correspondent tree either for non-contradiction or for third-excluded:

$$\begin{array}{c} (\mathbf{F}) \sim(p \wedge \sim p) \\ \text{- - - - - - - - - - - - -} \\ (\mathbf{F}) \sim p \\ (\mathbf{F}) \sim\sim p \\ (\mathbf{F}) p \end{array}$$

$$\begin{array}{c} (\mathbf{F}\ (p \vee \sim p) \\ \text{- - - - - - - - - - - - -} \\ (\mathbf{F}) p \\ (\mathbf{F}) \sim p \end{array}$$

The point is that for atomic p is no way to decompose further $\sim p$ and since the negation does not produce "switch of sides", we will not have $\mathbf{F}p$ and $\mathbf{T}p$ on the same branch.

- **Duality**: It seems that this negation has closed relations to the dual negation of GTS and IF-Logic. As pointed out in a personal communication by Tero Tulenheimo FDE-negation can be considered to be dual negation in the sense that it allows any formula to be formulated in an equivalent form in which negation appears at most in front of atoms. I will not delve into this issue very deeply here. However, I will discuss some points in paragraph **3** below.

Hintikka describes a tree-system that yields classical logic by adding two extra-closing rules[24]:

Namely

if it contains atomic formulae of the form $(\mathbf{T}) \sim p$ and $(\mathbf{T})p$.
if it contains atomic formulae of the form $(\mathbf{F}) \sim p$ and $(\mathbf{F})p$.

The first additional rule allows the validity of non-contradiction to be proved.
The second additional line allows the validity of third-excluded to be proved.

Some might wonder why Hintikka needs such an awkward way to obtain classical logic. Well, the point is that in the context of enquiry games it is important to distinguish the inferences that can be drawn from the premises alone (with the possible addition of information of the model), from those inferences that are drawn from the conclusion. In fact this allows implementing a kind of one-sided contraction. I will not come back to enquiry games in this paper.

[24] Hintikka, Halonen, and Mutanen [8, p. 50].

Interesting is that Hintikka and collaborators speak of such trees as being classical without any other comment. Similarly, Jeffrey's coupled trees were thought to capture—quite awkwardly—classical logic but in fact, as showed by Michael Dunn [3], were essentially trees for First Degree Entailment.

- In the present paper I will talk of *Hintikka trees** to refer to the trees described above without the addition of the two extra closure rules. In fact I will not come back to the classical version, so with the expression *Hintikka trees for enquiry games* I will always refer to the system that does not render classical logic.

In fact, Hintikka presented his tress in form of a sequent calculus and gave no explicit model theoretic semantics for it (since he took care to add classical rules).

We might retain from the sequent calculus the following point: FDE-negation does not *involve switch of sides*—meaning sides of a sequent or change of labels in a semantic tableaux system.

Let us now present Michael Dunn's relational semantics for FDE.

12.3.3 Michael Dunn's Relational Semantics for FDE

In the following we will make use of Michael Dunn's original paper and of Priest's description.[25]

The logic FDE is the core of a family of relevant logics developed by Anderson and Belnap, starting at the end of the 1950s. In FDE the conditional is defined away with the help of the disjunction (and negation). For higher degrees of entailment more has to be done.

By 1960 Michael Dunn discovered relational semantics for FDE as a result of his algebraic semantics. The idea is that instead of having truth-***functions***; truth-***relations*** are introduced, allowing a formula to be related to false (0) and true (1) or to neither of them. The fact that a formula α relates to 0 (relates to 0: $\alpha\mathsf{R}0$) does not mean that it is untrue, since the formula can also relate to 1 ($\alpha\mathsf{R}1$). The fact that a formula does not relate to 1 (it is untrue), does not mean that it relates to 0 (is false) since it might relate with neither. The recursive definitions are the expected ones:

$$(\alpha \wedge \beta)\mathsf{R}1 \text{ iff } \alpha\mathsf{R}1 \text{ and } \beta\mathsf{R}1$$
$$(\alpha \vee \beta)\mathsf{R}1 \text{ iff } \alpha\mathsf{R}1 \text{ or } \beta\mathsf{R}1$$
$$(\alpha \wedge \beta)\mathsf{R}0 \text{ iff } \alpha\mathsf{R}1 \text{ or } \beta\mathsf{R}1$$
$$(\alpha \vee \beta)\mathsf{R}0 \text{ iff } \alpha\mathsf{R}0 \text{ and } \beta\mathsf{R}0$$
$$\neg\alpha\mathsf{R}1 \text{ iff } \alpha\mathsf{R}0$$
$$\neg\alpha\mathsf{R}0 \text{ iff } \alpha\mathsf{R}1$$

It is like the classical definitions but it is neither assumed that the relation is exhaustive (a conjunction can be related neither with 1 nor with 0) nor is it assumed that it is exclusive (a conjunction can be related to both, 1 and 0).

[25] Dunn [3], Priest [18, pp. 139–41].

Semantic consequence is defined in the usual way in terms of truth-preservation, thus

$\Sigma \models \alpha$ iff for every model based on R, if $\beta\mathsf{R}1$, for all $\beta \in \Sigma$, then $\alpha\mathsf{R}1$.

Certainly, one could see relational semantics as defining truth-functional fourth-valued semantics for values true, false, both (b) and neither (n). Validity is then defined as preservation of the designed values true and b.

To relate the Hintikka-trees* for enquiry games with the relational semantics, the following yields the main idea of the procedure to read off counter-models from the open plays of a tree:

For every atomic formula p stated by the left (**T**-) side of the tableau (tree):

- if it is a positive literal ($\mathbf{T}p$) set the truth-value relation $p\mathsf{R}1$;
- if it is a negative literal ($\mathbf{T}{\sim}p$) set the truth-value relation $p\mathsf{R}0$

In the Appendix 12.4 it is shown that this is indeed the case proving soundness and completeness of Hintikka-trees* in relation to the relational semantics just described. Now this is all what we can really say. We have a semantics and a proof system that is sound and complete to it. I might be biased but the tableaux system is conceptually more elucidatory than the relational semantics. There is also a Routley-star semantics, but I will not delve into it in this paper. Be it like it be, let us see how the dialogical analysis look like.

12.3.4 A Dialogical Study of FDE-Negation

12.3.4.1 Switch of Choices: Is Duality-Negation a Tonk-Like Operator?

In the following we will render a dialogical semantics for the logic of FDE but only of the notion of negation involved.

A first dialogical approach to the semantics of negation in FDE-logic yields, as we will see below a particle-rule without switch of challenge-defence roles but switch of choices. Take disjunction, the dialogical semantics for the disjunction amounts to determine that the defender is the one who can choose which of the two disjuncts he will engage himself to prove. The negation of the disjunction will produce a switch of choices: the defender must defend the negation of one of the disjuncts chosen by the challenger.[26] The negation of FDE has been provided with several different semantics, namely a relational, a fourth-valued (described above) and a Routley-star semantics. All of them can be thought of in the dialogical framework as the negation of switch of choices. One might even argue that this

[26] It turns out that the negation described above is quite close to the dialogical approach to negation proposed by Laurent Keiff [11].

is what dual-negation is about in the dialogical framework (see 1.1 above). Dual negation is the basic negation of GTS, perhaps one way to distinguish between the analysis of negation in GTS and Dialogical Logic is how to understand the notion of switch.

We will first pursue this path but at the very end it the analysis is not deep enough. Moreover, such an approach contravenes against the third condition of dialogical harmony, namely the independence of the inner structure of the formula in the scope of the negation. But let us start, with the switch of choices study. This will split the particle rules in several sub-rules:

$\sim$	**Challenge**	**Defence**
X-$\sim(\alpha \vee \beta)$	**Y**-?$\sim\vee_1$ *or* **Y**-?$\sim\vee_2$ challenger chooses	**X**-!-$\sim\alpha$ *respectively* **X**-!-$\sim\beta$
X-$\sim(\alpha \wedge \beta)$	**Y**-?$\sim\wedge$	**X**-!-$\sim\alpha$ *or* **X**-!-$\sim\beta$ defender chooses

In fact, the rules for "$\sim$" yield the basis for first-degree entailment provided that double negation is assumed. Once more, there is no switch of burden of the defence:

$\sim\sim$	**Challenge**	**Defence**
X-$\sim\sim\alpha$	**Y**-?$\sim\sim$	**X**- α

We are not at the logic of first-entailment yet, but we are very near. What we need is to generalize the formal rule in order to include negative literals (negations of atomic formulae):

Formal rule for FDE (FR-FDE)
P cannot introduce literals: any literal (positive or not) must be uttered by **O** first. Literals cannot be challenged.

Validity is defined via the notion of winning strategy for **P** (as in standard logic).

The problem with this approach to negation is that it introduces a difference between literals and other complex formulae. Negative literals have no "local semantics" (particle rule), but only global semantics (structural rule). In fact one could formulate and ground a particle rule for negative literals. The point is that FDE-negation produces a change of choices and since there is no choice left make there is no defensive move possible. Let us call the resulting dialogic FDE^{L}:

Particle rule for FDE-negative literals:

Assertion	**Challenge**	**Defence**
X-$\sim p$	**Y**-?$\sim$	— No defence possible

Formal Rule for FDE^{L}:
P cannot introduce positive literals: any positive literal must be uttered by **O** first. **P** can challenge a negative literal iff the same negative literal (uttered by **P**) has already been already challenged by **O** before. Positive literals cannot be challenged.[27]

Examples: See Examples 2, 3, 4 Appendix 12.3.

12.3.4.2 Negation at the Structural Level of the FDE-Logic

The particle rule(s) formulated above apparently contravenes against the third condition of dialogical harmony. Indeed, the particle rule for this negation is not independent of the inner structure of the formula in which this logical constant occurs as a main operator. Let us then, leave the same particle-rule as the standard one, that is:

[27] To produce classical logic from FDE^{L} two further assumptions must be added, namely third-excluded and explosion.

Classical Assumptions for Dialogues With—Negation

1. Every dialogue contains as an initial hypothesis the **O** concession $p \bigvee \sim p$, for any literal p that occurs in the thesis.
2. If **O** utters a positive and a negative literal, **P** is allowed to utter an *arbitrary* atomic formula (if **P** is committed to utter it).

$\neg$	Challenge	Defence
X-$\neg\alpha$	**Y**-α	— No defence possible

But change the formal rule involving the logic of FDE

FDE-Negation Defined by Structural Rules

P cannot introduce literals: any literal (positive or not) must be uttered by **O** first.
P can utter the double negation of a positive literal if **O** uttered the correspondent negation-free literal before. This double negation utterance of **P** can not be challenged.
P can utter a positive literal if **O** uttered the double negation of the same literal before.
Literals cannot be challenged.

The latter formulation will render a negation that has the same meaning as the standard negation at the *local* level but a different meaning at the *global* level. From the dialogical point of view FDE- negation is basically switch of choices. The FDE semantics of negation stops where there are no more choices: that is at the literal level.

In both approaches to dual negation (namely, the one in which a new particle is introduce and the one in which new structural rules are added), negative literals have a special status and double negation is introduced as an axiom. One would expect that at least in the structural approach double negation results from some features linked to the notion of duality at stake and not from a substitution rule that implements it by force. We will explore this possibility in the following paragraph.

12.3.4.3 Dual Negation and Dual Dialogical-Contexts

One more general perspective on dual negation is to see this negation as relating to a switch to a parallel dialogical context in which the *dual* of the first formula is uttered.[28] The dual of a dual drives us back to the original dialogical context. This kind of retrieval operation is called *involution*. For the sake of simplicity I will assume that there are only two argumentative contexts, the initial one, with the label 0, and the context 0* containing the dual of the formula $\sim\varphi$ uttered at 0—though it is possible to set a more general frame where each dialogical context **j**

[28] This approach is approach is based on Routley's star-operator Routley [30].

has its star-"double" **j***—we will need this more general approach when we discuss conditionals.

The challenge to a dual-negation-operator commits the challenger to switch to a parallel dialogical context where he becomes the defender of the positive form (the dual) of the challenged negation. The point is that a challenge on a negation is an operation that carries the play to another context with a re-distribution of the roles of challenger and defender. It is different for example of the switch of contexts triggered by a challenge on a universal modality: the latter does not commit to the re-distribution of the roles in relation to the formula within the scope of the negation.

$\sim$	**Challenge**	**Defence**
i: **X**-$\sim\alpha$ (the negation has been uttered at context **i**=0 or **i**=0*)	**j**: **Y**- α (**Y** challenges the negation by uttering α at **j**, where: if **i**=0, then **j**=0* if **i**=0*, then **j**=0)	— No defence possible

This negation is equivalent to FDE-negation and thus does neither support the validity of third excluded nor of non-contradiction—the point is that we will never have in the same argumentative context $\sim\varphi$ and its challenge φ^*. That double negation holds (in a setting with classical structural rules) is a consequence of the assumption that the notion of duality involved has an associated involution operation.

The view that this approach to the meaning of dual negation sets is that negation is about switch between the roles of challenger and defender in dual contexts and this constitutes the rock bottom of a negation-operator. The switch of choices view developed is only one part of the story, namely the duality between disjunction and conjunction triggered by the dual negation.

Standard negation arise when the dual of a context is the contexts itself (the so-called self-dual) and then it reduces to switch in the roles of defender and challenger. But the latter switch is in fact characteristic of the conditional and is a partial switch—in a conditional the switch between defender and challenger happens only in relation to the head of the conditional.

12.3.4.4 The Conditional in Dual Contexts

The Conditional and the Switch of Players-Side

In the standard FDE-logic the conditional is simply defined as disjunction. Let us first recall what the standard definition of the dialogical conditional is. The point of dialogical meaning of the conditional is that it triggers a partial switch of the commitments to the defence. Indeed, after the challenge the defence of the conditional splits in two: the head has now to be defended by the challenger while the tail has to be defended by the original defender.

The strategy to simply combine dual negation with the standard particle rule for the conditional yields a very weak conditional since there will be almost no interaction between negation and the conditional besides the ones that FDE-logic allows for disjunctions. Indeed if, say, the tail of the conditional is negative, and it is even true, then, since this means that at the parallel context the positive part is uttered—and not at the context where the conditional is uttered, contraposition does not hold. What we need is to understand the conditional also as linked to a switch between dialogical contexts in general, with and without star. In other words we need to understand the conditional modally.

The formulae will be assumed to have the form **X**-α → β and label **i**, where **i** is a natural number and follows the labelling of dialogical contexts of a S5 system.

Assertion	Challenge	Defence
i: **X**-(α → β)	j: **Y**- α (the challenger utters the head of the conditional at an available context **j** of his choice, where **j** could be any label **n** and/or **n***)	**j**: **X**- β

It is crucial to point out that even if the conditional is understood here modally, and even if this modality involves contexts **i***, the possible switch to a parallel context triggered by this modality does not induce a redistribution of the roles of challenger and defender in relation **to the whole formula**. Indeed, the challenger **Y** is the player who utters the head of the conditional ***even*** at the (possibly new) context **j**.[29]

A total redistribution of the roles in a parallel context is obtained only by dual negation of a conditional. In our modal case, the challenge to the dual negation of a conditional uttered at **n** by player **X** will trigger a switch to a parallel dialogical context **n*** where **Y** utters the conditional in the scope of the negation. A further counterattack of **X** at **n*** will induce a play where **X** utters the head of the conditional and **Y** its tail.

[29] If we really would like to see negation as a conditional then a kind of modal minimal logic will result, where there are no contexts with stars.

$\to$	**Challenge**	**Defence**
i: **X**-$\sim(\alpha \to \beta)$	j: **Y**- $(\alpha \to \beta)$ (the challenger utters the conditional at the context **j**, where: if **i**=0, then **j**=0* if **i**=0*, then **j**=0)	— (no defence possible. Though **X** might launch a counterattack to the conditional by uttering the head of the conditional at a context of his choice)

- The rules for conjunction and disjunction are the standard ones except that these formulae carry labels of corresponding dialogical contexts, though neither the challenges nor the defences of the correspondent formulae induce switches to parallel contexts.

The resulting logic is called $\mathbf{K_4}$ by G. Priest.[30] The denomination is a bit confusing since it is in fact the result of adding an S5-strict implication to the framework of dual negation.

12.4 Conclusions

Local meaning is, according to the dialogical point of view, about rules that regulate specific actions, namely:

- player-independent utterances,
- how to raise a question in relation to an utterance (*local* challenge) and
- how to answer to a request (*local* defence).

Are dual negation and dual conditional tonk-like operators? No, they are logical constants and allow inconsistency but not triviality. It might be even defended that dual negation represents the core of the dialogical meaning of negation. Those actions that set the local meaning of negation and the conditional seem to be linked to a switch understood as:

a *partial* re-distribution of defender and challenger roles (dialogical conditional and accordingly defined negation)

a *total* re-distribution of defender and challenger roles carried out in parallel dialogical contexts. This kind of switch assumes an operation of involution and is linked to the further action of choosing sides in the context of conjunctions and disjunctions.

30 Priest [18, pp. 163–65]. See too Priest [19].

Perhaps conditional and negation are different special cases of the duality between defender and challenger roles—a notion of duality that constitutes one of the most basic elements of the dialogical approach to meaning. At this point we either start again or stop. I will take the second option for the moment . . .

Appendix 12.1 Note on Symmetric and Asymmetric Versions of the E-Rule

In the standard literature on dialogues, there is an asymmetric version of the intuitionistic rule, called E-rule since Felscher [4]. It's formulation is the following:

> **O** may react only upon the immediately preceding move of **P**.

This produces an asymmetry: **P** can challenge whatever formula but **O** can not.

In fact these devices yield intuitionistic logic if we add that only the most recent challenge (that has not been yet responded) can be defended.

In general to introduce the E-rule could be seen as jeopardizing the (for the dialogical approach crucial distinction) between the play and the strategic level. The rule has mainly a strategic motivation. According to my view the idea of ranks deployed in the text is more adequate. Indeed, the choice of the appropriate rank number to obtain validity is part of the strategic level. In fact if the ranks have been set such that Opponent has rank 1 and the Proponent rank two and the intuitionistic structural rule is in force, then the Proponent has a winning strategy for a propositional formula iff this formula it is valid in intuitionistic propositional logic. Now, to say that the Proponent has rank 2 is and Opponent rank 1 is very close to say that the Opponent can only react upon the immediately preceding.

According to this idea Rahman [20] proposed the following analysis of the role of the E-Rule in intuitionistic logic:

(1) The asymmetric E-Rule is based on strategic considerations, namely, the different roles in a strategy of the **P**- and the **O**-utterances.
(2) The symmetric E-Rule is based on meaning considerations, namely the specific local and global meaning of the conditional (and the negation as a special case), that allows locally to switch the roles of challenger and defender and might trigger globally defence delays.
(3) The asymmetric E-Rule yields a system of strategies that corresponds to Gentzen's Calculus of 1935, the symmetric E-rule is closer to Beth tableaux. Indeed, the tableaux corresponding to Gentzen [6] do not allow two formulae to occur at the right side (do not allow that two **P**-formulae occur at the same time in the same branch). Beth tableaux are more permissive.
(4) The asymmetric E-Rule allows straightforward proofs of some metamathematical properties of intuitionistic logic such as the interpolation theorem and the disjunctive property. For the latter see the following point.

(5) In Rahman's PHD it is shown how to prove the disjunctive property of intuitionistic logic with the asymmetric E-Rule. In his paper *Why Dialogical Logic?* [31] Rückert presents the argument with some detail. The point is that if we consider the distinction between the play and the strategic level then the proof of the disjunctive property can be carried out in the same way with symmetric or asymmetric rules (see Appendix 12.2). A more detailed presentation of the arguments involved have been published before by Rahman and Rückert in [24].

Appendix 12.2 The Disjunctive Property and the Symmetric Rule for Intuitionistic Logic

The presentation of the proof given below stems essentially from Rückert [31]. The point was raised in Rahman [20] and discussed at length in Rahman and Rückert [24]. Similar has been pointed out by Rahman [20] in relation to the existential property of intuitionistic logic—the argument works analogously.

The disjunctive property of intuitionistic logic says that $\alpha \vee \beta$ is valid iff α is valid or β is valid:

$$\models \alpha \vee \beta \models \alpha \text{ or } \models \beta$$

The proof from the right to the left is unproblematic. If **P** has a formal winning strategy for α or for β, it is evident that he has a winning strategy for $\alpha \vee \beta$:

P starts the dialogue by stating the thesis $\alpha \vee \beta$, and **O** attacks it with "?". In order to win **P** then just has to choose the disjunct he has a winning strategy for. If he has a winning strategy for α (or β) he has to choose α (or β respectively) to answer **O**'s attack.

On the level of plays the notion of meaning underlying intuitionistic logic is captured by the socalled intuitionistic structural rule that we called above the asymmetric E-Rule:

In any move, each player may attack a (complex) formula asserted by his partner

or he may defend himself against the last attack that has not yet been answered.

The crucial point of this rule with regard to the present argumentation is that **P** is not allowed to defend himself against an attack of **O** he has already answered, unless **O** renews his attack.

In our example this means that **P** is not allowed to defend himself again with β (or α) against the attack "?" on $\alpha \vee \beta$ if he has already defended himself with α (or β respectively).

Now we go to the level of strategies and recall the definition of validity in the dialogical approach: A formula is valid iff **P** has a formal winning strategy for it.

Thus, from the left to the right the meta-theorem of the disjunctive property says, that if **P** has a winning strategy for $\alpha \vee \beta$ he also has a winning strategy for at least one of the two disjuncts. And to have a winning strategy means for **P** that he is able to win the dialogue no matter how **O** plays.

If we look at the beginning of a dialogue with the thesis $\alpha \vee \beta$, it is clear that the first move of **O** has to be the attack "?" and that **P** has to reply α (or β) in the second move. Now, the dialogue continues with an argumentation about α (or β respectively) alone, if **O** does not renew his attack on $\alpha \vee \beta$. Consequently, if **P** has a winning strategy for $\alpha \vee \beta$, he must also have a winning strategy for at least one of the two disjuncts alone (independently of the hope that **O** might renew his challenge).
Quod erat demonstrandum.

Appendix 12.3 Examples

In the following examples, the outer columns indicate the numerical label of the move, the inner columns state the number of a move targeted by an attack. Expressions are not listed following the order of the moves, but writing the defence on the same line as the corresponding attack, thus showing when a round is closed. Recall, from the particle rules, that the sign "—" signalises that there is no defence against the attack on a negation.

For the sake of a simpler notation we will not record in the dialogue the rank choices but assume the uniform rank **O**: n=1 **P**: m=2:

Example 1: Classical and intuitionistic rules

In the following dialogue played with classical structural rules **P**' move 4 answers **O**'s challenge in move 1, since **P**, according to the classical rule, is allowed to defend (once more) himself from the challenge in move 1. **P** states his defence in move 4 though, actually **O** did not repeat his challenge—we signalise this fact by inscribing the not repeated challenge between square brackets.

O			**P**		
				$p \vee \neg p$	0
1	$?_{\vee}$	0		$\neg p$	2
3	P	2		—	
[1]	[$?_{\vee}$]	[0]		p	4

Classical rules. **P** wins.

In the dialogue displayed below about the same thesis as before, **O** wins according to the intuitionistic structural rules because, after the challenger's last attack in move 3, the intuitionist structural rule forbids **P** to defend himself (once more) from the challenge in move 1.

O				**P**	
				$p \vee \neg p$	0
1	$?_{\vee}$	0		$\neg p$	2
3	p	2		—	

Intuitionist rules. **O** wins.

Example 2: FDE^{L}-rules

O				**P**	
				$p \vee {\sim}p$	0
1	$?_{\vee}$	0		${\sim}p$	2
3	$?{\sim}$	2		—	

FDE^{L}-rules. **O** wins.

Ex. 3: FDE^{L}-rules

O				**P**	
				${\sim}$ (p$\wedge$${\sim}$p)	0
1	$?{\sim}\wedge$	0		${\sim}{\sim}p$	2
3	$?{\sim}{\sim}$	2			
[1]	$[?{\sim}\wedge]$	0		${\sim}p$	4
5	$?{\sim}$	4		—	

FDE^{L}-rules. **O** wins.

Ex. 4: FDE^{L}-rules

O				**P**	
H	${\sim}p$			${\sim}p \vee q$	0
1	$?\vee$	0		${\sim}p$	2
3	$?{\sim}$	2		—	
	—		H	$?{\sim}$	4

FDE-rules. **P** wins.

Appendix 12.4 Soundness and Completeness of Hintikka-Trees* for Enquiry Games in Relation to M. Dunn's Relational Semantics for FDE

The following proofs are standard and there is no claim of originality here beside the remark that Hintikka trees* describe FDE. In fact the proofs are a variation of Dunn's prove [3] in relation to Jeffrey's coupled trees.[31]

12.4.1 Soundness

The main job is done by the following definition 1 the rest is mechanical

Definition 1 (DS1):

Let us consider a set $\boldsymbol{S}$ of signed formulae such as $\mathbf{T}\ \alpha \wedge \beta$, ${}_{i.n}\mathbf{F}\ \alpha \vee \beta$, occurring in a branch B. We say that $\boldsymbol{S}$ is faithful in the relational model $\boldsymbol{M}$ defined by means of the relation R described before if there is a mapping $\boldsymbol{f}$ such that:

(a) If $\mathbf{T}\alpha$ is in $\boldsymbol{S}$, then the mapping yields $\alpha\mathsf{R}1$ in $\boldsymbol{M}$
(b) If $\mathbf{F}\alpha$ is in $\boldsymbol{S}$, then the mapping is such that it is not the case that $\alpha\mathsf{R}1$ is in $\boldsymbol{M}$ (that is, α might be related either to 0 or to neither but not with 1)

- We say that a branch of a tree is faithful if the set of signed formulae on it is faithful in some model.
- We say that a tree is faithful if some branch of it is faithful

Soundness Lemma 1 (SL1):

A closed tree is not faithful.

Proof:

- Suppose that we have a tree closed and faithful.
- Since it is faithful, some branch of it is. Let $\boldsymbol{S}$ be the set of formulae on that branch and let it be faithful in the model $\boldsymbol{M}$ by means of the mapping $\boldsymbol{f}$.
- Since the tree is closed then for some atomic formula α we must have $\mathbf{T}\alpha$ and $\mathbf{F}\alpha$. But then $\alpha\mathsf{R}1$ must be in $\boldsymbol{M}$ but this is impossible since the mapping of $\mathbf{F}\alpha$ will prevent this to happen.

Soundness Lemma 2 (SL2):

If (a section of) a tree is faithful and a branch of that (section of the) tree is extended by tree-rules, the result is another faithful (section of) a tree.

[31] See also Priest [18, pp. 152–59].

(Obviously this assumes that the formula that triggers the extension is not atomic).

Proof:

Let $\mathcal{T}$ be a faithful tree and let **B** be the branch that is extended.

The proof requires several steps. We begin with two main steps:

By hypothesis at least one branch is faithful, now this branch could be **B** or could be **B***.

(I) if the faithful branch is **B*** the extension of **B** will leave **B*** unchanged, thus after the particle rule has been applied to **B**, $\mathcal{T}$ will still be faithful (because **B*** is).
(II) if the faithful branch is **B** and it faithful in the model ***M*** the proof is by cases.

That is, by the consideration of all the ways to extend the branch **B** by the application of the corresponding tree-rule to a labelled and signed formula at the end of that branch. Namely by the application of a **T**-signed conditional-rule and a **F**-signed conditional rule etc.

Let us work with $\mathbf{F}(\alpha \wedge \beta)$. Since the model is faithful to the branch, it is not the case that $(\alpha \wedge \beta)\mathsf{R}1$ is in ***M***.

Hence, either it is not the case that $\alpha\mathsf{R}1$ or it is not the case that $\beta\mathsf{R}1$. Thus, at least one of the branches will be faithful and hence the extended tree will be faithful too.

Soundness Theorem:

If there is Hintikka-tree-proof of α, α is valid in the relational semantics for FDE.

Proof:

Assume that there is tree-proof of α, but α is not valid. We show that from this a contradiction follows.

Since there is a (Hintikka-)tree-proof of α there is a closed tree $\mathcal{T}$ that starts with $\mathbf{F}\alpha$. Thus, the first section of $\mathcal{T}$ is $\mathcal{T}_0$ that consists in the thesis $\mathbf{F}\alpha$. The following sections of $\mathcal{T}$ are constructed by extending $\mathcal{T}_0$.

Since α is not valid, there is some model M at which α is not true. Let assume this model and an adequate mapping such that $\{\mathbf{F}\alpha\}$ is faithful to M. Thus $\mathcal{T}_0$ is faithful, since the set of formulae on its only branch is faithful.

Since $\mathcal{T}_0$ is faithful by lemma SL2 so is any tree $\mathcal{T}$ we get that starts with $\mathcal{T}_0$ and results by extending $\mathcal{T}_0$.

It follows that $\mathcal{T}$ is faithful.

$\mathcal{T}_0$ is closed by hypothesis, and this is impossible by SL1.

12.4.2 Completeness of Hintikka-Trees for Enquiry Games in Relation to M. Dunn's Relational Semantics for FDE Worked Out Trees*

Let us say that a Hintikka-tree* (for enquiry games) has been *worked out* if all appropriate tree-rule applications have been made.

What we need is a systematic method for constructing a tree that ensures that if we start a tree for a given formula, we can always produce a worked out a tree. There are many options available the following, as we will discuss below. Now, systematic worked out trees do not contribute to no insight it is only a mechanical procedure:

1. First stage: thesis
2. Any other stage:

   2.1 pick up the leftmost open branch
   2.2 pick up a formula that is neither an atomic formula and do the following going from top to bottom:

      2.2.1 If it is a formula of the form **F**-$\sim\varphi$, simply apply the appropriate rule to it. In such a way that the resulting subformula will be added to the end of each open branch on which this negation occurs only if it does not occur there before. Similar for **T**-$\sim\varphi$,
      2.2.2 If the formula is a **T**-conjunction or a **F**-disjunction add both of the corresponding subformulae to the end of each open branch on which the conjunction/disjunction occurs only if it does not occur there before. Similar for a **F**-conditional
      2.2.3 If the formula is a **T**-conjunction **or a** **F**-disjunction split the end of each open branch on which the conjunction/disjunction occurs and add the corresponding subformulae only if it does not occur there before. Similar for a **T**-conditional

   (Tick any formula that has been subject of the application of a rule)

After all this has been done, do the same for the second from the left and so on.

Comments

- Notice that systematic trees avoid repetitions. In fact it is a systematic formulation of the non-repetition rule.

Definition 1: Construction of a relational model from a branch **(CD1)):**

Let us consider a systematically worked out (Hintikka-)tree* with the open branch **B**. We show how to construct a relational model ***M*** in which **B** is faithful for α atomic and occurring in **B**.

1. **T**α iff αR1 in ***M***
2. **T**$\sim\alpha$ iff αR0 in ***M***

Completeness Lemma (CL):

THESIS:

For *each* formula Φ (atomic or not) on the open branch **B** of a given worked out tree we can determine a relational model ***M*** in the following way:

(a) If **T**α occurs in **B**, then αR1 is in ***M***
(b) If **F**α occurs in **B**, then it is not the case that αR1 is in ***M***

(that is, α might relate to 0 or to neither but not to 1)

(c) If **T**∼α occurs in **B**, then αR0 is in ***M***
(d) If **F**∼α occurs in **B**, then it is not the case that αR0 is in ***M***

(that is, α might relate to 1 or to neither but not to 0)

Proof:

By induction on the complexity of the formula Φ.
The property at stake in our case is the one described in the thesis

BASE CASE: Assume Φ is the atomic formula α

(1) If **T**α occurs on **B**, by CD1 we have αR1 in ***M*** as required by (a).
(2) If **F**α occurs on **B**, then, since the branch is open, **T**α does not occur. Hence, by CD1, it is not the case we have αR1 in ***M*** as required by (b).
(3) If **T**∼α occurs in **B**, by CD1 we have αR0 in ***M*** as required by (c).
(4) If **F**∼α occurs in **B**, then, since the branch is open, **T**∼α does not occur. Hence, by CD1, it is not the case we have ∼αR1 in ***M***, but then, by the relational semantics for the negation it is not the case that αR0 as required by (d).

INDUCTION CLAUSE

Assume (induction hypothesis) that we know the result for formulae simpler than Φ.

Let us start with the case Φ is **T**(α ∧ β). If **T**(α ∧ β)occurs on **B**, since the tree has been worked out, the following formulae occur on **B**, too:

Tα
Tβ

Since, these are simpler than Φ, by induction hypothesis we will have in ***M*** that:

αR1 ***M***. and
αR1***M***.

Thus, as required, in M we have

$$(\alpha \wedge \beta)\mathsf{R}1$$

Let us take **F**$(\alpha \wedge \beta)$. If **F**$(\alpha \wedge \beta)$ occurs on **B**, since the tree has been worked out, one of the following formulae occur on **B**, too:

$$\mathbf{F}\alpha \text{ or } \mathbf{F}\beta$$

Then by induction hypothesis, we have that

> it is not the case that $\alpha\mathsf{R}1$ or it is not the case that $\beta\mathsf{R}1$ is in ***M***. Thus, it is not the case that $(\alpha \wedge \beta)\mathsf{R}1$ is in ***M***.

Negated conjunction and disjunction are proven similarly.

Suppose that **T**$\sim\alpha$ occurs on **B**. Since we know the case of for **T**$\sim\alpha$ whenα is atomic, we have that, as required, $\alpha\mathsf{R}0$ is in ***M***. Similarly for **F**$\sim\alpha$.

Suppose that **T(F)**$\sim\sim\alpha$ occurs on **B**. Then, **T(F)**α occurs on **B**. Hence by induction hypothesis, $\alpha\mathsf{R}1$ is (is not) in ***M***.

Completeness Theorem:

> THESIS:
>
> If Φ is valid in a relational semantics, then there is a proof using rules of Hintikka's trees for Enquiry games

We prove the contrapositive:

> If there no proof in the tree-system at stake, Φ is not valid.

If we start the proof with **P**Φ, and work out a systematic tree, it will not close (since by assumption of the contraposition there is no proof for **P**Φ,).

Let us pick up the open branch **B** of the tree. With help of the CD1 we can create a relational for which the thesis CL holds. In particular, if **P**Φ' occurs on B, then it is not the case that $\Phi\mathsf{R}1$ is in ***M***. But **P**Φ occurs on **B**, since it is the signed formula at root of the tree, so it is on every branch. Thus there is a models such that Φ does not relate to 1, so it is not valid in the relations semantics for FDE.

Quod erat demonstrandum

References

1. Clerbout, N. 2011. "Dialogical Games for First-Order Logic and Tableaux." Paper presented at the *14e Congrès de Logique, Méthodologie et Philosophie des Sciences,* 19–26. Juillet, 2011, Nancy (France).

2. Clerbout, N., L. Keiff, and S. Rahman. 2009. “Dialogues and Natural Deduction.” In *Acts of Knowledge, History, Philosophy, Logic*, edited by G. Primiero and S. Rahman, chap. 4. London: College Publications.
3. Dunn, M. 1976. “Intuitive Semantics for First-Degree-Entailments and Coupled-Trees.” *Philosophical Studies* 29:149–68.
4. Felscher, W. 1985. “Dialogues as a Foundation for Intuitionistic Logic.” In *Handbook of Philosophical Logic*, edited by D. Gabbay and F. Guenthner, vol. 3, 341–72. Dordrecht: Kluwer.
5. Fiutek, V., H. Rückert, and S. Rahman. 2010. “A Dialogical Semantics for Bonanno’s System of Belief Revision.” In *Constructions*, edited by P. Bour et alii, 315–34. London: College Publications.
6. Gentzen, G. 1935. “Untersuchungen Ueber das Logische Schliessen.” *Mathematische Zeitschrift* 39:176–210.
7. Hintikka, J. 1999. *Inquiry as Inquiry: A Logic of Scientific Discovery*. Dordrecht: Kluwer.
8. Hintikka, J., I. Halonen, and A. Mutanen. 1999. “Interrogative Logic as a General Theory of Reasoning.” In Hintikka [1999], 47–90.
9. Keiff, L. 2004a. “Heuristique formelle et logiques modales non normales.” *Philosophia Scientiae* 8–2:39–59.
10. Keiff, L. 2004b. “Introduction à la dialogique modale et hybride.” *Philosophia Scientiae* 8–2:89–105.
11. Keiff, L. 2007. “Approches dynamiques à l’argumentation formelle.” PhD thesis, Université de Lille, Lille.
12. Keiff, L. 2009. *Dialogical Logic, Entry in the Stanford Encyclopaedia of Philosophy.* http://plato.stanford.edu/entries/logic-dialogical/.
13. Keiff, L., and S. Rahman. 2010. “La Dialectique entre logique et rhétorique.” *Revue de Métaphysique et Morale* Avril-June 2010 2:149–78.
14. Lorenz, K. 2001. “Basic Objectives of Dialogue Logic in Historical Perspective.” *Synthese* 127:255–63.
15. Lorenzen, P. 1955. *Einführung in die operative Logik und Mathematik*. Berlin/Göttingen/Heidelberg: Springer.
16. Lorenzen P., and K. Lorenz. 1978. *Dialogische Logik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
17. Prawitz, D. 1979. “Proofs and the Meaning and Completeness of the Logical Constants.” In *Essays on Mathematical and Philosophical Logic*, edited by J. Hintikka, I. Niiniluoto, and E. Saarinen, 25–40. Dordrecht: Reidel.
18. Priest, G. 2001. *An Introduction to Non-Classical Logic*. Cambridge, MA: Cambridge University Press.
19. Priest, G. 2011. “Realism, Antirealism and Consistency.” In present volume.
20. Rahman, S. 1993. *Über Dialogue, Protologische Kategorien und andere Seltenheiten*. Frankfurt/Paris/New York, NY: P. Lang.
21. Rahman, S. 2009. “A Non Normal Logic for a Wonderful World and More.” In *The Age of Alternative Logics*, edited by J. van Benthem et alia., 311–34. chez Dordrecht: Kluwer-Springer.
22. Rahman, S., N. Clerbout, and L. Keiff. 2009. “Dialogues and Natural Deduction.” In *Acts of Knowledge, History, Philosophy, Logic*, edited by G. Primiero, 301–36. London: College Publications.
23. Rahman, S., and L. Keiff. 2004. “On How to Be a Dialogician.” In *Logic, Thought and Action*, edited by D. Vanderveken, 359–408. Dordrecht: Kluwer.
24. Rahman, S., and H. Rückert. 1998. “Die pragmatischen Sinn und Geltungskriterien der Dialogischen Logik beim Beweis des Adjunktionsatzes.” *Philosophia Scientiae*, 1998–1999 3/3:145–70.
25. Rahman, S., and H. Rückert, eds. 2001. *New Perspectives in Dialogical Logic*. Special volume of *Synthese* 127.

26. Rahman, S., and T. Tulenheimo. 2009. “From Games to Dialogues and Back: Towards a General Frame for Validity.” In *Games: Unifying Logic, Language and Philosophy*, edited by O. Majer, A.-V. Pietarinen, and T. Tulenheimo, part III. Dordrecht: Springer.
27. Read, S. 2008. “Harmony and Modality.” In *Dialogues, Logics and Other Strange Things: Essays in Honour of Shahid Rahman*, edited by C. Dégremont, L. Kieff, and H. Rückert, 285–303. London: College Publications.
28. Read, S. 2010. “General Elimination Harmony and the Meaning of the Logical Constants.” *Journal of Philosophical Logic* 39:557–76.
29. Redmond, J., and M. Fontaine. 2011. *How to Play Dialogues. An Introduction to Dialogical Logic*. London: College Publications, London.
30. Routley, R., and V. Routley. 1972. “The Semantics of First Degree Entailment.” *Noûs* 6: 335–95.
31. Rückert, H. 2001. “Why Dialogical Logic?” In *Essays on Non-Classical Logic*, edited by H. Wansing, 165–85. New Jersey, NJ/London: World Scientific.
32. Rückert, H. 2007. *Dialogues as a Dynamic Framework for Logic*. PhD-thesis, Leyden, 2007. http://openaccess.leidenuniv.nl/dspace/bitstream/1887/12099/1/R%C3%BCckert_PhD_Dialogues_neu.pdf.
33. Rückert, H. 2011. “The Conception of Validity in Dialogical Logic.” Talk at the Workshop Proofs and Dialogues, Tübingen, Organized by the Wilhelm-Schickard Institut für Informatik (Universität Tübingen), February 25–27, 2011.
34. Schröder-Heister, P. 2008. “Lorenzen’s Operative Justification of Intuitionistic Logic.” In *One Hundred Years of Intuitionism (1907–2007)*, edited by M. van Atten, P. Boldini, M. Bourdeau, and G. Heinzmann. Basel: Birkhäuser.
35. Smullyan, R. 1968. *First-Order Logic*. Heidelberg: Springer.
36. Sundholm, B. G. 1983a. “Constructions, Proofs and the Meaning of the Logical Constants.” *Journal of Philosophical Logic* 12:151–72.
37. Sundholm, B. G. 1983b. “Systems of Deduction’, chapter I:2.” In *Handbook of Philosophical Logic*, edited by D. Gabbay and F. Guenthner, vol. I. Dordrecht: Reidel.
38. Sundholm, B. G. 2010. “Proofs as Acts and Proofs as Objects: Some Questions for Dag Prawitz.” *Theoria* 64(2–3):187–216.
39. Tulenheimo, T. 2010. O*n the Dialogical Approach to Semantics*. Talk at the Workshop Amsterdam, CA/Lille: *Dialogues and Games: Historical Roots and Contemporary Models*, February 8–9, 2010, Lille. http://www.tulenheimo.webs.com/talks.html.

Chapter 13
Necessary Truth and Proof

Stephen Read

13.1 Truthmaker Realism

The realist believes that truth is in no way dependent on our ability to detect it, so there can be verification-transcendent truths, truths which may forever elude our attempts to discover their truth. In what then, does their truth consist, if it is so divorced from our investigative powers? What the realist has resorted to over many years is the language of truth-making, now formalised into a theory of truth, Truthmaker Realism. At its heart lies the principle Truthmaker, which appears in various, mostly inequivalent, forms. The one I find plausible is the principle of the **Supervenience of Truth on So-Being**:

ST Truth supervenes on how things are: there can be no difference in truth without a difference in how things are.

This is distinct from two stronger principles, that of the **Supervenience of Truth on Being**, that differences in truth require differences in what exists, and the yet stronger principle, **Truthmaker**, that whatever is true, something makes it true.

The main objection to this last principle, **Truthmaker**, concerns negative existentials, e.g., 'There are no unicorns' or 'Vulcan does not exist'. It is not that there is something which makes them true, rather, they are true because something (unicorns, Vulcan) does not exist. The objection to the **Supervenience of Truth on Being** is similar. Take, e.g., the fact that I am sitting down. This is true not because I exist, or that a standing me does not exist, but because of how I am, that I am sitting and not standing. Truth supervenes on how things are, not on what there is. Truth supervenes on so-being rather than being, using the terminology of Meinong's Principle of Independence. Nonetheless, in some cases, truth will supervene on what there is, will even be made true by what there is, as **Truthmaker** says. Take essence, for example. If a being's nature is essential to it, then what supervenes on its essence,

Reprinted with kind permission of the editors of Kriterion

S. Read (✉)
University of St Andrews, St Andrews, Scotland, UK
e-mail: slr@st-andrews.ac.uk

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_13,

or nature, will supervene on it too, since it could not exist without its essence, that is, without being as it is.

I want to defend the twofold thesis that necessary truths require truthmakers, and that what make necessary truths true, are proofs. In insisting that every truth, including necessities, needs a ground for its truth, I am following Leibniz, who wrote in his *Monadology* (§§31–2):

> Our reasonings are based on two great principles, the *principle of contradiction* ... and the *principle of sufficient reason*, by virtue of which we consider that we can find no existent fact, no true assertion, without there being a sufficient reason why it is thus and not otherwise.

Truth supervenes on so-being, and even necessary truths require explanation why they are true. Indeed, they require more. They require explanation why they are necessarily true. That is supplied by the proof, which if it exists, exists of necessity, and entails immediately that its conclusion is true and necessary.

Why do necessary truths require truthmakers? It is often claimed that only contingent truths need truthmakers. Necessary truths cannot fail to be true, and so there is no explanatory need for anything to make them true. But this is circular: if they must be true, then they will indeed be true, but why must they be true? Simply calling them necessary truths does not make them true. Two thoughts are often adduced:

1. Necessary truths are entailed by all truths, so whatever makes any truth true makes all necessary truths true too. Hence necessary truths are made true, not by their own special truthmakers, but by all truthmakers, and indeed, the fact that they are made true by all truthmakers, not just some, is what makes them necessarily true.
   This argument depends on two contentious premises:

   NAQ Necessary truths are entailed by any proposition whatever, and
   ET Truthmaking is closed under entailment, that is, any proposition is made true by the truthmakers of any proposition which entails it.

   I reject **NAQ** for logical reasons, along with its flipside, **EFQ** (that an impossible proposition entails any proposition whatever); and I believe **ET** needs qualification, as below.
2. Necessary truths are empty of content, a doctrine left over in positivism from Wittgenstein's *Tractatus* even when its rationale there had been rejected. Hence there are no facts corresponding to necessary truths, which are purely analytic and so true solely in virtue of the meaning of the constituent words. Thus they require no truthmaker.
   One reason to reject this line of thought goes along with the rejection of **NAQ**: not all necessary truths are equivalent (which they would be, given **NAQ**) and so they cannot all be empty of content.

Given the rejection of 2, and of **NAQ**, one needs to reconsider **ET**. For **ET** also suggests that entailment itself has no content and no truthmaker. But if 'α entails β' is made true by, say, s and α is made true by t, then β must be true (since α and 'α

entails β' are true). But what makes β true? It cannot just be t (as decreed by **ET**) but some combination of s and t. Some notation: I write 'α entails β' as $\alpha \Rightarrow \beta$, and 's makes α true' as $\mathsf{s} \models \alpha$ (i.e., s *forces* α). Then we can endorse a revised entailment principle:

RET If $\mathsf{s} \models \alpha$ and $\alpha \Rightarrow \beta$ then $f(\mathsf{s}) \models \beta$,

where $f(\mathsf{s})$ is some function of s. Of course, until something is said about the function f, this really says no more than that if α is true and α entails β, then β is true. What will give it substantive content is to relate f to the proof that makes $\alpha \Rightarrow \beta$ true.

An obvious objection to the claim that all necessary truths are established by proof appeals to the arguments of Kripke and Putnam. Their arguments appear to show that all true identities, including many a posteriori identities, are necessarily true. But a posteriori truths do not admit of proof, or they would be a priori.

There are two possible responses. First, suppose Kripke and Putnam are right. Nonetheless, the role of proof here is not to establish the truth of true identities but their necessity. What Kripke, following Barcan Marcus and others did, was to show that if $a = b$ is true then it is necessarily true. There is a general proof-scheme into which one slots '$a = b$' and whose conclusion is that $a = b$ is necessarily true. But this conclusion inherits from its empirical premise its own a posteriori status.

It's not clear if this solution is coherent. A better approach for the defender of "proof-maker realism" may be to reject Kripke and Putnam's arguments. They depend on too simplistic a semantics for modality and necessity (and too gullible a use of thought-experiments). The point is a general one about the utility of logic. Presented with a philosophical puzzle one may turn to logic to model it. The logic may then deliver an unpalatable result. Rather than accept the result, one may query the logic (in this case, the modal semantics), until a reflective equilibrium is reached. In the present case, that may only come once a semantics for modality has been formulated which admits contingent identity.[1]

In proposing proofs as truthmakers, I am not defending a form of conventionalism. But where traditional conventionalism falls to two objections, the involvement of proofs in establishing their truth forestalls them. The objections are first, Quine's observation [15] that the conventionalist account is circular, appearing to explain necessity as what necessarily followed from certain conventions, and secondly, that it fails to recognise the obvious difference between trivial necessities (e.g., 'Bachelors are unmarried') and deep necessities (e.g., Fermat's Last Theorem). The distinction between trivial proofs and deep proofs immediately speaks to the second of these objections. The first objection is dealt with by the autonomy of logic and mathematics. Logical and mathematical concepts are autonomous in that their meaning is given by the rules for their use. They are self-justifying. Permitting certain inferences regarding a logical or mathematical term, inferences which a

[1] For such a semantics for contingent identity, see, e.g., [14, chap. 17].

convention or decision that the term should have a certain logical or mathematical meaning encapsulate, has ineluctable consequences for what follows from assertions containing the term. These consequences conform to the principle that one can legitimately extract from an assertion as much, but no more than, one must have correctly to make it.[2]

There are three further objections to the idea that proofs could act as truthmakers, however, First, a famous result of Kurt Gödel's may be invoked, where he appeared to show that truth outstrips proof, in that he showed that in any mathematical system of sufficient strength, there are true assertions which are not provable if the system is consistent. Secondly, it may be objected that the role of proofs is not ontological, to make certain truths true (and necessary), but epistemological, to demonstrate to ourselves that they are true. Finally, it may be objected that the project is circular, in that a proof only demonstrates that its conclusion must be accepted if its axioms are, and so the idea of proofs as truthmakers is regressive, explaining the truth of the conclusion at the expense of requiring demonstration of the truth of the axioms.

The third objection falls to the same inferentialist response as did Quine's objection above, in explaining how the meaning of logical and mathematical terms is encapsulated in the valid inferences in which they appear. The inference rules are self-justifying or autonomous. The first and second objections are dealt with seriatim, in §§2 and 3.

13.2 Incompleteness

Hilbert's programme was a response to the nineteenth century crisis in the foundations of mathematics. The nineteenth century opened with the seminal contributions of Cauchy and others towards getting clear about the foundations of analysis, as it had been developed from Newton's and Leibniz' beginnings. In particular, their elucidations needed to deal with the blatant inconsistency involved in, for example, dividing by a number which in the limit was zero. The end of the century saw the creation of modern set theory by Cantor and others, as a general language and formal theory which would, it was believed, at long last provide the firm and consistent framework for nineteenth century analysis. But the century ended in worse chaos than that in which it started, the vague suspicions of inconsistency at the beginning being replaced by formally proved contradictions at the close, such as Cantor's and Burali-Forti's paradoxes. The new century opened with the discovery of the most famous of the formal contradictions, that of Zermelo and Russell, of the set of all sets which are not members of themselves.

Hilbert's reaction, like others such as Russell's, was a retrenchment, an attempt to salvage as much as possible of classical analysis. But some careful restriction of mathematical procedures was clearly needed, and Hilbert's diagnosis, like Brouwer's, was that the problem lay in unfettered use of the infinite. What was

[2] The inferentialist answer is developed further in [17, 19].

needed was some finitistic restriction on mathematics together with a proof that when finitistic constraints were (in careful ways) lifted, the result would remain consistent. It was taken for granted—or supported by a traditional inductive proof—that finitistic methods could not induce contradiction; the lesson from the preceding centuries was that the problem arose only when finite methods were extended to deal with infinite classes. Behind this methodological approach lay an ontological interpretation. Talk of the finite was of the real; talk of the infinite should be treated "as if" it were true and referred to actual entities, but no reliance should be placed on this, and no derivation or use of the infinite was complete until all statements involved were again open to a finitistic interpretation. This ontological conception backed up the methodology. For orthodox philosophy, embodied for Hilbert in the critical philosophy of Kant, believes that the real cannot be contradictory. It is the ideal elements, Hilbert's "as if" or Kant's "phenomena" (appearances), which not only can be but are contradictory, as in the paralogisms. Hence arose the need in Hilbert's programme for proofs of the consistency—and finitistic proof, at that, of course—of the use of the ideal. Proof of conservative extension then shows that any inconsistency is restricted to the ideal elements.

Arithmetic was a seemingly straightforward case in point. Peano's postulates have no finite models. Every number has a successor, if the successors are equal the numbers are equal, and 0 is not a successor. So there can be no greatest number, and no circles. Hence all models are infinite. Thus there was an urgent need to show that arithmetic, formulated clearly and exactly in some such way as Peano's and Dedekind's, was consistent, and the expectation was that a proof would be readily forthcoming. As is well known, Gödel dashed these hopes, and with them, any likelihood that Hilbert's programme could trace a way out of the crisis. For Gödel showed first, that any sufficiently strong system of arithmetic such as Peano's, if consistent, is incomplete, in leaving some arithmetical truth unproven; and more appositely, as a corollary, that if consistent, its consistency cannot be demonstrated within arithmetic, and so a fortiori cannot be demonstrated in any finitistic subsystem.

The relevance of Gödel's result for us is that it appears to show that no system of proof is adequate to the role of truthmakers, for there are arithmetical statements which though true are, relative to any particular system of proof, unprovable in that system. Of course, any particular truth can be proved in some system, e.g., one in which it is an axiom. But no single system can establish the truth of all arithmetical truths, and so it cannot be their proofs (in that system) which make them true. The mistake in this reasoning is to accept Gödel's account of what constitutes a proof as relevant to our concerns. It is not. Gödel's concept of proof, which has become the norm, is of an array of formulae whose conformity to a set of rules is decidable, that is, one which can be checked in a finite time as being a proof or not. This accords with the epistemological needs of Hilbert's programme, and probably with other epistemological reservations. But we are not constrained by epistemology. It is completely plausible that some necessary truths, including arithmetical truths, are true in virtue of proofs which it is, for whatever reason, impossible for us to formulate. If this seems surprising, it is surprisingly easy to give an example.

There are, in Peano arithmetic formulated on the basis of first-order predicate logic, two ways to prove a universal proposition. Suppose the logical basis is given in natural deduction. Then besides the ∀I-rule:

$$\frac{A(u)}{\forall x A(x)}\,\forall\mathrm{I}$$

(where 'u' is not free in any assumption on which the premise depends), whose conclusion has the form of a universal proposition, there is the rule of Mathematical Induction (Ind):

$$\frac{A(0) \quad A(u) \to A(u+1)}{\forall x A(x)}\,\mathrm{Ind}$$

(where again, 'u' is not free in the assumptions on which the second premise depends), whose conclusion has the same form. However, in both cases, establishing the conclusion requires following a uniform method for every number, in the first to show $A(u)$, where u is arbitrary, in the second to show $A(u) \to A(u+1)$, again where u is arbitrary. One might conjecture, in the face of Gödel's result, that the incompleteness of arithmetic lies in the requirement of uniformity, required for the decidability of the Gödelian notion of proof. For suppose we introduce a rule with infinitely many premises, often termed the rule of infinite induction or ω-rule[3]:

$$\frac{A(0) \quad A(1) \quad A(2) \quad \ldots\ldots}{\forall x A(x)}\,\omega\text{-rule}$$

where there is a premise of the form $A(n)$ for every natural number n. Two things should be immediately obvious: first, we cannot now check in a finite time whether a proposed array, containing at least one instance of the ω-rule, conforms to the rules, for there are infinitely many sub-proofs to be checked. Hence, provability is no longer decidable. Thus, arithmetic with the ω-rule is not subject to Gödel's theorem. The proof predicate is not primitive recursive. Secondly, it is fairly obviously complete, since the side-condition on the ω-rule exactly matches the truth-condition for 'all', so $\forall x A(x)$ is true (and provable) iff each instance $A(n)$ is true (and provable).

The unprovable formula in Gödel's theorem is just such a universal formula, saying that it has no proof: it has the form $\forall x(T(x) \to \neg\exists y \mathsf{Prov}(y, x))$, where '$T$' holds uniquely of this formula, and $\mathsf{Prov}(y, x)$ says that y is (the Gödel number of) a proof of (the wff with Gödel number) x.[4] So too, is Goldbach's Conjecture, that every even number (greater than 2) is the sum of two primes. (Hence Gödel [11, p. 305] referred to his undecidable sentences as of "Goldbach form".) Goldbach's Conjecture has resisted proof for a long time, and it may be that its proof cannot proceed in the usual way by a uniform method, but requires demonstration in a

3 [5, §14], [21, pp. 258–61].

4 Clearly, in the presence of infinite proofs, the notion of Gödel number will need to be generalized.

different way for each natural number (or at least, for an infinite partition of the numbers) that it conforms to the Conjecture. Any such particular proof is obviously beyond us in practice, but *not* beyond our comprehension—we could not survey the indefinitely many different methods (if we could, there would be some way to treat them uniformly and finitely, contrary to hypothesis). But the general method of proof, via an infinitary rule of proof, the ω-rule, is easily seen to be both complete and, in itself, consistent (i.e., consistent relative to the system without it).

Although first-order arithmetic with the ω-rule is complete, second-order arithmetic with the same rule (i.e., based on second-order logic) is not, for there are second-order arithmetical truths (whose existence is no counterexample to the completeness of first-order ω-arithmetic) which cannot be proven in second-order ω-arithmetic. The reason is that there are uncountably many subsets, that is, properties, of natural numbers. But the problem here is not with proofs of the first-order quantified formula, $\forall x A(x)$, but the second-order formula $\forall F A(F)$. For this to be true, it is not sufficient that it be true of some ω-sequence of properties F, e.g., the recursive predicates (or sets), or the definable properties. For these constitute a small island of countability in a sea of uncountability. But now language seems to constrain us, for, as normally conceived, there are only countably many predicate expressions. It is not enough that $A(F)$ be true for all predicates F. We need it true for all properties F. But uncountably many of these cannot be expressed in a countable language.[5]

The moral is not to be defeated, or intimidated by infinity. Indeed, this was Cantor's lesson, when he tamed the transfinite. Even with our confinement to finite modes of expression, we can gain an understanding and comprehension of indefinite orders of infinity. Consider, e.g., Paris and Harrington's proofs of arithmetical truths not provable in Peano arithmetic. Goodstein's theorem, for example, which claims that the Goodstein function eventually takes the value 0 for a sufficiently large, but finite argument, cannot be proved in first-order Peano arithmetic (of the kind subject to Gödel's theorem), but its proof using the assignment of transfinite indices (i.e., by transfinite induction) is concise and transparent. Note once again, that our comprehension of these proofs is not the point. The essential point to realise is that there are these proofs, many as yet undiscovered, and perhaps incapable in the end of discovery (not least, since there are infinitely many of them), whose existence is licensed by the meaning of the terms involved.

13.3 Anti-realism

The idea of verification-transcendent truth has been challenged many times over the past 100 years, however. Can there be propositions whose truth we may be incapable of establishing? In particular, the challenge has been refined into a dilemma: of any

[5] The assumption that languages are countable can be challenged. For example, infinite decimals are expressions (expressions using base-10 notation), though no (non-recurring) infinite decimal can ever be written down. Nonetheless, there are uncountably many such decimals, as the usual diagonalization argument shows.

proposition, either we are capable in principle of recognising its truth or falsity or if not, then we cannot have conferred on the relevant expression any clear meaning, that is, it does not express any proposition. If an expression has a clear meaning, then there are circumstances whose obtaining is necessary and sufficient for the truth of the proposition expressed and which we can recognise to obtain when they do, and not when they do not.

The main thesis which Dummett uses to support these claims is the manifestation challenge, that every aspect of meaning so conceived must be capable of manifestation. The argument for this conclusion relies on the consideration that meaning can and must be learned. Language is a social activity which is transmitted from generation to generation. Somehow, the child or language-learner comes to understand the meaning of the terms of the language, and indeed, those terms have no meaning other than that given to them by the practice of linguistic communication. Hence there can be no element of meaning which is not exhibited in some linguistic behaviour and which the language-learner can come to appreciate. All aspects of meaning must be capable of being manifested and acquired from participation in the practice.

The meaning of a proposition determines when it has been correctly uttered and when not. In particular, its meaning is such that, uttered in certain circumstances it is true, uttered in others it is false. In order to be capable of manifestation, those truth-conditions must be ones we can recognise to obtain or not to obtain. For if we could not do so, we could not make manifest how its truth depended on those circumstances and so could not articulate its meaning in terms of those truth-conditions. Hence the conditions of truth of any proposition with a determinate meaning must be capable, in principle, of being recognised as verifying or falsifying it. There can be no verification-transcendent propositions.

To the extent to which this argument is compelling we are here confronted by a paradox. For it seems that there are clear cases of propositions we understand yet which we also realise we could never verify. Dummett gives an example himself: 'A city will never be built here'.[6] It is straightforward enough to recognise that the proposition has been falsified: finding a city there. But to verify it, we would need to complete an infinite task of checking throughout eternity that no city had been built, and however late we left our check, that would verify only up to that time that no city had been built, leaving countless ages of subsequent history when one might appear. The example is reminiscent of the situation we noted might obtain with Goldbach's Conjecture, which might require checking separately and independently of each even number that it was the sum of two primes. Indeed, the situation seems even more likely in the empirical case. It seems implausible that there be a general reason, valid for all times, why a city will not be built—though, of course, there could be such a reason, such as the ineradicable presence of toxic heavy metals, or the absence of an adequate water supply. Even so, to be assured that a city will never

[6] [8, p. 16].

be built, means ensuring these obstacles will never be overcome, and that seems to require an open-ended check which would remain forever uncompleted.

How should this paradox be resolved? The claim is not that we cannot understand such a proposition as ‘A city will never be built here’. The claim is that the requirement that we be capable of manifesting a grasp of its meaning entails that we should draw back from a realist avowal of bivalence for it—that it must be either true or false—and so that we should manifest our understanding not in terms of a grasp of when it is true or false (for we have no justification for a belief that it must be one or the other), but in terms of when we are justified in asserting it and when are justified in denying it. Our human limitations necessitate that we could not envisage a circumstance obtaining whose recognition we would take as showing the proposition true, for that would require infinite knowledge of the absence of a city at all times. Hence our understanding and its manifestation must relate to what grounds we would accept as justifying its assertion or denial. ‘Never’ gets its meaning from quantification over finite and surveyable domains. Extrapolating it to the supposed indefinite future yields a proposition we could never be in a position to assert—a verification-transcendent one—and so one whose truth-conditions we cannot conceive. Consequently, we are not entitled to claim it is either true or false, and must reject a realist interpretation of there being a state of affairs whose existence makes it true or false.

The fault with all such anti-realist arguments is that they systematically underestimate our conceptual powers. Dummett [8, p. 17] concedes that although we cannot ourselves check on the possible truth of ‘A city will never be built here’, we can conceive of a being (he calls it ‘God’) who could verify it, by an infinite check. Our conception of the truth of the proposition is what such a being would have verified by verifying each successive instance—‘A city has not been built here yet’, ‘A city has still not been built here’, and so on into the indefinite future. If we can conceive of such a being, then we can comprehend what such a being would have verified and of which it had such infinite knowledge. To be sure, there are anxieties, precisely those which prompted Hilbert’s programme and the intuitionistic philosophy of mathematics, about the coherence of such an extrapolation of our finite powers. Yet there had always been a scepticism about the coherence of the notion of the infinite. Cantor’s bold step was to propose a careful and systematic treatment of the infinite on a par with the finite.

When the notorious antinomies such as Cantor’s and Russell’s were discovered, a natural reaction was to continue such scepticism about the notion of the infinite, and blame them on Cantor’s inception. But a careful diagnosis of the antinomies reveals that they offer “old wine in new bottles”. Russell’s paradox of the set of all sets which are not members of themselves has an immediate analogy which Russell recognised in the paradox of heterologicality, of the adjective which applies to all adjectives which do not apply to themselves, hence to the property of not applying to itself, that is, of not being true-of-itself, and so relates back to the infamous Liar paradox, of the proposition which says of itself that it is not true. That Russell (and Zermelo) discovered the paradox by reflection on the proof of Cantor’s theorem does not show that something must be inherently wrong with Cantor’s theorem itself.

Rather, Cantor's world of the transfinite offered new possibilities for old paradoxes to arise again. What is needed is a satisfactory diagnosis of those paradoxes, not a hasty and universal ban on all talk of the infinite, any more than an analogous ban on self-reference in the face of the semantic paradoxes.

What lies at the heart of the realist's confidence in his position is, we noted, a belief that reality cannot be incoherent or inconsistent. Only our descriptions, beliefs and theories can. Clearly, something was wrong with the foundations proposed for late nineteenth century mathematics. But the anti-realist reaction, Hilbertian, Brouwerian or whatever, which jettisons belief in an underpinning reality in favour of revised forms of description, cannot be the right one. There cannot be anything wrong with infinite collections in themselves. The error must lie in our descriptions of them. The descriptions created by Cantor and others needed revision; but the collections themselves are as real as ever, and so are either as described or not.

Consider, once again, 'A city will never be built here'. According to (**ST**), either there is something whose being built makes it false or whose absence makes it true. Clearly, that thing is a city. If a city is built here, the proposition will be false; so if it is true, whatever would have falsified it must not have existed or be going to exist. This is no argument for bivalence and realism, of course. Rather, it is an affirmation, and explanation, of the realist's belief in there being a truth-condition for the proposition.

Dummett's other famous (purported) counterexample to bivalence is also unconvincing, but for a different reason. Consider Jones, he says, who was never placed in circumstances which might have established whether or not he was brave, and is now, sadly, dead. Was Jones brave? Dummett's proposal is that the sense of 'brave' is dispositional, such that 'Jones was brave' means 'If Jones had encountered danger, he would have acted bravely'. That may be, but he claims that 'Jones was not brave' means 'If Jones had encountered danger, he would not have acted bravely'.[7] Although Stalnaker's semantics for counterfactuals validates Conditional Excluded Middle, this is not acceptable, as Lewis [13, pp. 79–82] observed. Jones' bravery is as good a counterexample as others: if Jones was not brave, he might nonetheless have behaved by chance as if he were. The correct analysis of 'Jones was not brave' is 'If Jones had encountered danger, he might not have acted bravely'.[8] So analysed, Jones is definitely either brave or not, in that either he would have acted thus or he might not have. Dummett's example only fails to satisfy Excluded Middle because he analyses it wrongly. If we can show that Jones might not have acted appropriately in the relevant circumstances, we have shown that it is false that he would have acted so, and hence false that he was brave. Thus, whatever the limitations on our now showing it, either Jones was brave or he was not.

---

[7] [8, p. 15]; cf. [9, pp. 342, 347].

[8] In symbols, Dummett expresses 'Jones was brave or Jones was not brave' as $(\alpha \mathrel{\Box\!\!\rightarrow} \beta) \vee (\alpha \mathrel{\Box\!\!\rightarrow} \neg\beta)$, where '$\mathrel{\Box\!\!\rightarrow}$' is the subjunctive conditional, 'If it were α it would be β'. The correct analysis is $(\alpha \mathrel{\Box\!\!\rightarrow} \beta) \vee (\alpha \mathrel{\Diamond\!\!\rightarrow} \neg\beta)$, where $\alpha \mathrel{\Diamond\!\!\rightarrow} \beta$ is equivalent to $\neg(\alpha \mathrel{\Box\!\!\rightarrow} \neg\beta)$, that is, it is not the case that if it were α it would not be β, i.e., if it were α it might be β.

But we cannot leave the example there. Perhaps Lewis can: the behaviour of Jones' counterparts acts for Lewis as a truthmaker of the modal proposition about Jones. I eschew belief in the reality of other worlds and the existence of their denizens. Fortunately, I am not presently being tortured. So there is no state of affairs of my being tortured, not even a non-actual such state of affairs. Without being actual, there can be no unity to constitute a state of affairs.[9] Nor is there any counterpart of me who is being (non-actually) tortured. (What an awful thought, that I escape torture only at his expense.) So such counterparts cannot act as truthmakers. What, then, does make 'Jones was brave' true or false? It is the fact that Jones was, or was not brave, his actual mental qualities. It is Jones' make-up which determines his behaviour, and so determines in the fiction, the behaviour of his counterparts—how he would act in such and such circumstances, not vice versa. It was disbelief in the bare truth of dispositional analyses which provided the initial spark for talk of truth-makers. C. B. Martin could not accept that phenomenalists or logical behaviourists had offered a satisfactory analysis of perception or belief or whatever if it ended in a conditional concerning possible sense-data or actions. What could make such conditionals concerning unperceived sense-data or unperformed actions true if all there are, are actual sense-data and actions? Unless these theories came up with elements of a kind they sought to avoid (such as Russell's sensibilia or categorical materialist bases), they were not only incomplete but incompletable. All truths must have a truthmaker, whether that truthmaker is a categorical basis (for Armstrong) or a power (for Martin) or a thing *qua* truthmaker (for Lewis). Or, by **ST**, at the very least, there is some property which Jones has or lacks which grounds his disposition to behave.

Dummett's comments show that, if he is right, we can accept the demand for truthmakers and still resist realism. But his grounds for resistance are mistaken. It is not our faith in the reality of the infinite which creates the problem. It merely allows existing problems to reappear in dramatic and virulent form. There are truths whose meaning and truth-conditions we can understand and yet whose truth (or falsity) we may be forever incapable of recognising. In particular, our understanding of mathematical notions allows us to conceive of there being proofs whose existence may lie beyond our ability to grasp or survey them. Nonetheless, if they do exist, then they establish the truth of their conclusions of necessity.

13.4 Logical Pluralism

Although intuitionism began as a philosophy of mathematics, a revisionary programme aiming for a fresh approach to the foundational crisis in late nineteenth century mathematics, intuitionistic logic has of late had a new lease of life in theoretical computer science. It is one of the prime examples of a new breed of logical relativism, what has come recently to be called "logical pluralism". The idea

---

[9] For an elaboration of this argument, see [18].

is that there are many logics, some suited for one purpose, others for another. A leading example is so-called linear logic. Its inventor, Girard, frequently remarked that "linear logic is not just another exotic logic". What he meant was that linear logic was an all-embracing approach to computational issues, in which the actual logic was but a tool in the analysis of computation. This logic had in fact been considered, at least in a fragmentary way, by Church [6] in his "weak theory of implication". Once again, Church did not advocate his weak theory as the ultimate truth about implication. Rather, it was a useful tool with which to explore the differential effects of certain implicational assumptions. The weak theory encapsulated everything which (Church thought) was uncontentious about implication (primarily prefixing, suffixing and permutation), and any other assumptions (above all, contraction and weakening) could then be separately considered. So too with linear logic. Basic assumptions are built in (though in subsequent variants, even these have been revised) and the particular effects of non-linearity, namely, contraction and weakening (i.e., multiple and zero or vacuous uses of assumptions) can be separately studied through the modal ("exponential") operators.

Linear logic started as a methodology, not a philosophy of logic. It proves its usefulness as a tool for the study of computation, but some now claim linear proofs give an insight into the notion of valid inference.[10] The reverse was the case with intuitionism. Originally conceived as a philosophical response to the foundational questions concerning mathematics, its recent popularity has been methodological, once again, as a tool for the study of computation. The interest is in what can be achieved—that is, computed—with the finite, but theoretically unlimited, resources of mechanical procedures. Rarely is this coupled to any claim that non-computable functions are suspect or non-constructive results not to be trusted. Rather, the attraction of intuitionistic, or more generally, constructive methods is their usefulness in studying what can, in principle, be implemented on a computer.

At the heart of this methodology lies the so-called Curry-Howard correspondence, or Curry-Howard isomorphism. By this correspondence, the formulae of propositional intuitionistic logic are seen to match the types of function terms in a λ-calculus, such that '&' matches Cartesian product, '$\rightarrow$' matches functional application, '$\vee$' matches disjoint union (or direct sum) and '$\perp$' (absurdity) matches the empty type. A function term is well-formed just when its type is (intuitionistically) provable, and the proof articulates how the value of the term is computed.

Such a constructivist interpretation of intuitionistic logic does not constitute logical pluralism, however. For intuitionistic provability is taken to show computability, not validity. Logical pluralism is better illustrated by the claim, for example, that classical logic is valid in finite domains, intuitionistic logic valid in infinite domains, and perhaps, relevance, or some other paraconsistent logic, valid in inconsistent domains (e.g., databases). This doctrine is refuted by an argument of Langford's and Quine's.[11] They mistakenly took the argument as a defence of classical logic. It cannot be that, at least, not without some supplementary consideration in favour of distinctively classical principles. What it does show is that logic is not relative to

[10] See the radical anti-realism of Jacques Dubucs: e.g., [7, p. 214].

[11] [12, p. 582], [16, chap. 6].

other purposes. It does not make sense to say that an argument is valid for purposes X but not valid for purposes Y. The reason is that truth is not relative, as Plato showed. For either truth is relative, in which case relativism is false for me (or Plato), or it is not relative. Either way, the doctrine of the relativity of truth is false. So too for any suggestion that validity is relative. For to say that an argument is valid is to say that the truth of the premises guarantees the truth of the conclusion, that is, that the argument preserves truth. So if the argument were valid for purposes X, it would preserve truth. Hence it preserves truth for purposes Y, since truth is independent of what purposes one has.

Suppose, for example, that it is claimed that classical logic is valid in finite domains, but only intuitionistic logic is valid in infinite domains. Then there will be counterexamples to classical principles in infinite domains. But if there are counterexamples to a principle, it is not valid. So classical logic is not valid, regardless of what domain is in question. Or suppose it is said that classical logic is valid in consistent situations, but only a paraconsistent logic is valid in inconsistent situations. Then there are counterexamples to classical principles, and so those principles are not valid *tout court*.

Constructivism is acceptable, therefore, as a methodology for the study of finitary procedures, but it must be rejected as an account of validity and truth. For there are verification-transcendent truths, as we noted in §3. Moreover, what should be resisted is the all-pervasive finitary interpretation of terms such as 'proof', 'construction' and 'procedure'. It is nowadays almost universal to find these terms given a finitary meaning. We noted this sense of 'proof' in §2. The very title 'constructivism' exhibits the phenomenon, for 'constructive procedures' is now taken to mean 'procedures determining an outcome in a finite time'. But traditionally, many constructions were non-terminating, or as one would now say, "non-constructive". The terminology fortunately lives on in many mathematical textbooks. Consider the famous "construction" of the reals from sets or cuts of rationals. Stewart and Tall, in their classic work on *The Foundations of Mathematics*, write:

> we must solve the problem of constructing a complete ordered field $\mathcal{R}$ starting from $\mathcal{N}$, ... First the integers $\mathcal{Z}$ are constructed from $\mathcal{N}$, and the rationals $\mathcal{Q}$ from $\mathcal{Z}$. To construct $\mathcal{R}$ from $\mathcal{Q}$ is a more taxing operation .... [20, p. 173]

Ebbinghaus, in his book on *Numbers*, writes:

> Quite apart from its use in the definition of real numbers, the CANTOR construction with fundamental sequences has turned out to be the most fruitful. [10, p. 40]

Such constructions are infinitary, and entirely "non-constructive". The Dedekind cut construction, for example, identifies the irrational number $\sqrt{2}$ with the infinite (completed) totality of all rationals whose square is less than 2.

What I have done elsewhere [17] is to use some of the insights of constructive (i.e., finitary) proof theory in giving a realist account of meaning in terms of proofs. The proofs in question are arrays of formulae according with carefully specified rules, but with no constraint that it be possible to check, in a finite time, whether those rules have been obeyed. That is, whether a particular array is a proof may itself be verification-transcendent. Nonetheless, I claim, the concept of proof articulated there is perfectly comprehensible and coherent. But I do not propose to embark on that articulation here. I want to close by considering two further aspects of modality, namely, possibility and contingency.

13.5 Contingency

My main topic is necessity, and what makes necessary truths true. But there is another side to that coin, and it would be a mistake to suppose that one can develop a theory of the one independently of the other. In particular, recall that I rejected the suggestion that what makes it true that, for example, I might undergo torture is that some counterpart of me in another possible world does do so. Such possible worlds are a mere *façon de parler* and cannot serve as truthmakers. But the question then arises, what is the truthmaker of this proposition? What does make possibilities possible?

Some possibilities are possible because they are actual, of course. If I am in agony, then it is clearly possible for me to be in agony. For $\ulcorner p \urcorner$ entails that $\ulcorner p \urcorner$ is possible, for all p. So by (RET), if $s \models p$ then $\alpha(s) \models \Diamond p$, where α is the one-step inference of $\Diamond$I.[12] But suppose $\ulcorner p \urcorner$ is false, but nonetheless, contingently so. Then $\ulcorner p \urcorner$ is still possible. What makes it true that $\ulcorner p \urcorner$ is possible, and indeed, that it is contingent? It cannot be the truthmaker for $\ulcorner p \urcorner$, since $\ulcorner p \urcorner$ has no truthmaker—*ex hypothesi* it is false.

David Armstrong claims that what makes $\ulcorner \Diamond p \urcorner$ true in this case is whatever makes $\ulcorner \neg p \urcorner$ true. His argument is this[13]: suppose $\ulcorner p \urcorner$ is false and contingent. Then $\ulcorner \neg p \urcorner$ is true and contingent also. But given $\ulcorner \neg p \urcorner$ and given that it is contingent, the truth of $\ulcorner \Diamond p \urcorner$ is entailed. He draws back from endorsing (ET) in full generality, but he claims that it seems to hold in a wide variety of cases, such as here. Armstrong [3] spells out the argument in more detail in his contribution to the *Festschrift* for Hugh Mellor. Suppose $T \models p$, that is, T makes p true. Let us represent 'it is contingent that p' by '$\mathcal{C}p$'. Then $T \models \mathcal{C}p$. For, since $\ulcorner p \urcorner$ is contingent, so too is the existence of T. "Could the contingency of T lie outside T?", he asks. "It does not seem possible. It cannot be a relation that T has to something beyond itself. So T is the truthmaker for the proposition 'p is contingent' ".[14] Armstrong's reasoning can be formalized like this: suppose $T \models p$ and $\ulcorner p \urcorner$ is contingent. Then $T \models \mathcal{C}p$. He continues (loc.cit.):

Suppose	(1) $T \models p$	
Then	(2) $T \models \mathcal{C}p$	by the above reasoning
So	(3) $T \models p \,\&\, \mathcal{C}p$	
But	(4) $p \,\&\, \mathcal{C}p \Rightarrow \Diamond\neg p$	
So	(5) $T \models \Diamond\neg p$	by his (restricted) Entailment principle

[12] See, e.g., [19]. Note that a one-step inference is a function, from one proof to another.

[13] See [2, pp. 154–5]; [3, p. 15]; [4, §7.2].

[14] [3, p. 15].

The final steps of this argument are trivial. $\ulcorner\mathcal{C}p\urcorner$ means $\ulcorner\Diamond p \mathbin{\&} \Diamond\neg p\urcorner$, so clearly if $T \models \mathcal{C}p$ then $T \models \Diamond\neg p$, or at least, something very close to T does so, if we prefer (RET) to (ET) or some such version. The real puzzle is the step from line (1) to line (2), from $T \models p$ to $T \models \mathcal{C}p$, provided $\ulcorner p\urcorner$ is contingent. Armstrong's stated concern is to avoid, if he can, postulating a special categorical property of contingency *in re*, which he hopes the above argument will allow him to do. But what is the justification for line (2), rather than its motivation? Armstrong [4, p. 84] appeals to **ET**. He claims that

(*) if p is contingent, then p entails $\Diamond\neg p$.

So, since $T \models p$, $T \models \Diamond\neg p$ by **ET**. But one cannot appeal to **ET** here, for $\ulcorner p\urcorner$ does not entail $\ulcorner\mathcal{C}p\urcorner$, even when $\ulcorner p\urcorner$ is contingent. Indeed, no wff of the form $p \Rightarrow \mathcal{C}p$ is valid in S5, for atomic p. In fact, note that $\ulcorner\mathcal{C}p\urcorner$ is itself never contingent, for $\mathcal{C}p \Rightarrow \Box\mathcal{C}p$ (given S5-principles). If $\ulcorner p\urcorner$ is contingent, then it is necessary that $\ulcorner p\urcorner$ is contingent. But it is a basic tenet of relevant logic that the necessary does not follow from the contingent, and indeed by a result of Coffa's, contingencies do not relevantly entail necessitives (that is, wffs of the form $\Box\alpha$) unless they already contain them as a part.[15] Whatever makes $\ulcorner\mathcal{C}p\urcorner$ true must make it necessarily true. So suppose $\ulcorner p\urcorner$ is contingent, and that the proposition asserting the existence of what makes it true does not contain any necessitives essentially. Then it cannnot make $\ulcorner\mathcal{C}p\urcorner$ true, for by Coffa's result, the proposition asserting its existence cannot entail $\ulcorner\mathcal{C}p\urcorner$, since $\ulcorner\mathcal{C}p\urcorner$ is a necessitive. Whatever makes (contingent) $\ulcorner p\urcorner$ true must exist only contingently, but what makes $\ulcorner\mathcal{C}p\urcorner$ true must exist of necessity. So far from being a reason for it, (1) is inconsistent with (2). What makes $\ulcorner p\urcorner$ true, where $\ulcorner p\urcorner$ is contingent, cannot make $\ulcorner\mathcal{C}p\urcorner$ true.

This ironically shows that Armstrong's subsequent defence of (*) collapses. Armstrong writes:

> Given the attractive S5 modal logic, if p is contingent, it is a necessary truth that it is contingent. This may help to quell any doubts we may have about step [(*)] in the argument. [4, pp. 84–85]

Quite the contrary. It is the fact that (*) has a necessarily true conclusion which shows that it must fail.

Reverting to our original case, we are left with the original puzzle: if $\ulcorner p\urcorner$ is false but contingent, what makes it possible that p? The answer lies in a remark of Pierre Bayle's, cited in Leibniz' *Theodicy* (§173): "the possible is whatever does not contain a contradiction." $\ulcorner\Diamond p\urcorner$ is equivalent to $\ulcorner\neg\Box\neg p\urcorner$, and that is true, by (**ST**), if there is no proof of $\ulcorner\neg p\urcorner$. According to (**ST**), $\ulcorner\neg p\urcorner$ can be true simply by default, that is, in the absence of a truthmaker for $\ulcorner p\urcorner$. But that is not enough for the truth of $\ulcorner\Box\neg p\urcorner$, that is, to make it necessarily false that p. For that we need something to ensure that $\ulcorner p\urcorner$ is false, and as we have seen, necessary truths require proofs to make them true. But what would a proof of $\ulcorner\neg p\urcorner$ be? We can learn here, as so often,

[15] See [1, §§22.2.1–2]. Armstrong [4, p. 10] concedes that the entailment in **ET** must be relevant.

from the constructivists. Intuitionistic logic is often developed on the basis of the connectives '&', '$\vee$', '$\to$' and '$\perp$', and '$\neg p$' is defined as '$p \to \perp$'. Indeed, this is sometimes referred to as an intuitionistic or constructivist definition of negation. But its constructive character is given by imbuing '$\to$' with that character. Realist negation can also be taken to identify '$\neg p$' with '$p \to \perp$', provided the theory of '$\to$' is suitably strong and realist, and for a suitable choice of '$\perp$'.[16]

Hence, what makes it possible that I might be in insufferable agony is that nothing rules it out, that there is no proof that shows that the assumption that I am in agony is absurd. Fortunately, I am not. But there is nothing absurd in the suggestion that I might be. Hence it is possible.[17]

In general, necessary truths require proofs to make them true and non-necessary truths require their absence. Some of those proofs will be simple and straightforward; others can be more complex than the human imagination can conceive. Nonetheless, what the existence of such proofs depends on are the basic inferences which encapsulate the meanings of the terms and operations involved.

Acknowledgments This paper was presented at the conference '(Anti-)Realisms, Logic and Metaphysics': Nancy, 28 June–1 July 2006 and first published in *Kriterion* 50 (2010), 47–67.

References

1. Anderson, A., and N. D. Belnap. 1975. *Entailment: The Logic of Relevance and Necessity,* vol. 1. Princeton, NJ: Princeton UP.
2. Armstrong, D. M. 2000. "Difficult Cases in the Theory of Truthmaking." *Monist* 83:150–60.
3. Armstrong, D. M. 2002. "Truthmakers for Modal Truths." In *Real Metaphysics: Essays in Honour of D. H. Mellor*, edited by H. Lillehammer and G. Rodriguez-Pereyra, 12–24. London: Routledge.
4. Armstrong, D. M. 2004. *Truth and Truthmakers*. Cambridge, MA: Cambridge University Press.
5. Carnap, R. 1937. *The Logical Syntax of Language*. London: Routledge. A. Smeaton, Eng. Tr.
6. Church, A. 1951. "The Weak Theory of Implication." In *Kontrolliertes Denken: Untersuchungen zum Logikkalkül und der Logik der Einzelwissenschaften*, edited by A. Wilhelmy, A. Menne, and H. Angstl, 22–37. Freiburg-im-Breisgau: Karl Alber.
7. Dubucs, J. 2002. "Feasibility in Logic." *Synthese* 132:213–37.
8. Dummett, M. 1978. "Truth." In idem, *Truth and Other Enigmas*, 1–24. London: Duckworth.
9. Dummett, M. 1991. *Logical Basis of Metaphysics*. London: Duckworth.
10. Ebbinghaus, H. D. 1991. *Numbers*. New York, NY: Springer.
11. Gödel, K. 1972. "Some Remarks on the Undecidability Results." In *Collected Works*, edited by S. Feferman et al., vol. II, 305–06. Oxford: Oxford University Press, 1990.
12. Langford, C. H. 1928. "Concerning Logical Principles." *Bulletin of the American Mathematical Society* 34:573–82.
13. Lewis, D. 1973. *Counterfactuals*. Oxford: Blackwell.

[16] See [17].

[17] Note that nothing I have said above commits me to the S5 principle $\Diamond p \to \Box\Diamond p$. If one were committed to it, one would need a proof that there is no proof of $\neg p$. More plausibly, a theory of necessity as proof will reject $\Diamond p \to \Box\Diamond p$ and endorse an S4 theory of modality.

14. Priest, G. 2008. *An Introduction to Non-Classical Logic: From If to Is* (2nd Edition). Cambridge, MA: Cambridge University Press.
15. Quine, W. V. O. 1936. “Truth by Convention.” In *Philosophical Essays for A. N. Whitehead*, edited by O. H. Lee, 90–124. New York, NY: Longmans.
16. Quine, W. V. O. 1970. *Philosophy of Logic*. Englewood Cliffs, NJ: Prentice-Hall.
17. Read, S. 2000. “Harmony and Autonomy in Classical Logic.” *Journal of Philosophical Logic* 29:123–54.
18. Read, S. 2005. “The Unity of the Fact.” *Philosophy* 80:317–42.
19. Read, S. 2008. “Harmony and Modality.” In *Dialogues, Logics and Other Strange Things: Essays in Honour of Shahid Rahman*, edited by C. Dégremont, L. Kieff, and H. Rückert, 285–303. London: College Publications.
20. Stewart, I., and D. Tall. 1977. *The Foundations of Mathematics*. Oxford: Oxford University Press.
21. Tarski, A. 1956. “The Concept of Truth in Formalized Languages.” In *Logic, Semantics, Metamathematics*, edited by J. H. Woodger, 152–278. Oxford: Clarendon Press.

Chapter 14
Anti-realist Classical Logic and Realist Mathematics

Greg Restall

14.1 Introduction

My aim in this paper is to apply a semantically anti-realist understanding of (classical) logical consequence, and to then use the change of perspective from the semantically realist concern of *truth-in-a-model* to the semantically anti-realist analysis in terms of *propriety-of-assertion* (or *denial*) as a position from which to view the philosophy of mathematics. The result is not so much a new position in the metaphysics or epistemology of mathematics, but instead a fresh perspective on traditional positions.

Let us start with logical consequence.

14.2 Logic

The relationship between logic and mathematics is remarkably close. The rise of classical logic, in the work of Frege, Russell, Gödel and Tarski, arose not so much from a desire to give a uniform account of judgement, to treat problems of quantification in natural languages, to treat vagueness—target was *mathematics*. From the development of the calculus and its rigorisation, to the paradoxes of set theory, the aim was to clarify, to make explicit the forms of deduction valid in mathematical reasoning.[1] How we think of mathematics and how we think of logic are intertwined. In this paper, a refigured view of logic will bring along with it a reconfigured position in the philosophy of mathematics.

[1] For an enlightening historical account of the development of the mathematical sciences and logic, along with with the rise of the 'Semantic Tradition,' see Coffa's *The Semantic Tradition from Kant to Carnap* [5].

G. Restall (✉)
School of Historical and Philosophical Studies, Melbourne, VIC, Australia
e-mail: greg@consequently.org

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_14,

14.2.1 Logic Without Vocabulary

The logic of Frege, Russell, Gödel and Tarski is deductive logic. It is not only deductive logic, it is *classical* deductive logic. Our task here will not be to define a new relation of logical consequence, but to picture this relation in a distinctive way. We will reframe the central notion of logical consequence in terms of coherence. Coherence is a normative notion, making explicit a particular kind of *mistake* that one can make, or can avoid making, in a discourse.

The motivation, which will suffice to introduce the concept, is that coherence is a kind of virtue that a position in a discourse might have. A position—involving a collection of assertions and denials—is *coherent* when those assertions and denials hang together, when they are consistent. However, we will not think of coherence as defined in terms of possible truth or truth-in-an-interpretation. Instead, we will consider the upshot of taking coherence as a *starting* point for our analysis, as opposed to a notion defined in other terms.

Taking coherence as a starting point does not mean that there is nothing more we can say about it. On the contrary, it can be argued that coherence must satisfy certain criteria: there are *norms* of coherence [22].

NORMS: The norms are straightforward to state, given some individuation of the *content* of the assertions and denials of the discourse in question. We will call a collection of assertions and denials a POSITION. A position $[\Gamma : \Delta]$ is a pair of (finite) sets, Γ of things *asserted* and Δ of things *denied*. Positions are evaluated for *coherence*. Such an evaluation must satisfy the following three conditions:

IDENTITY: $[A : A]$ is not coherent.

WEAKENING: If $[\Gamma, A : \Delta]$ or $[\Gamma : A, \Delta]$ is coherent, then is $[\Gamma : \Delta]$ coherent too.

STRENGTHENING: If $[\Gamma : \Delta]$ is coherent, then either $[\Gamma, A : \Delta]$ or $[\Gamma : A, \Delta]$ is coherent too.

There are *many* different relations satisfying these norms. Perhaps the coherence relation on our target vocabulary is the smallest relation satisfying these conditions (the propositional contents are totally independent of one another; think of atomic statements in some formal language), or the relation is richer than this (think of the content of judgements in some particular language; we may say that the position [this is red : this is coloured] is incoherent).

These Norms Are Sufficient for Logical Consequence

If we take a discourse to be governed by a coherence relation satisfying these norms, then we thereby may evaluate it with respect to deductive validity. A logical consequence relation is *definable* in terms of coherence. For example, if we take a position $[\Gamma : A, \Delta]$ to be incoherent then, given that an agent has asserted (implicitly or explicitly) Γ and denied (implicitly or explicitly) Δ and is *coherent*, then the *only* coherent option available concerning A is to assert it. Once the question has come up, its answer is implicit. A is now undeniable, it *follows* from what has already been said. A kind of consequence is implicit in the notion of coherence.

These Norms Are Necessary for Logical Consequence

We can make the connection in the other direction too. Consider the kind of grip a *deductive argument* from A to B ought have on a discourse. It is too much to think that an assertion of A need be followed by an assertion of B, or that anyone who accepts A must accept B. What we require is that the assertion of A is not to be combined with the denial of B—that a position in which A is asserted and B is denied is defective.

So, suppose that we take an argument from premises to a conclusion[2] to have the normative force of rendering 'defective' a position in which the premises are asserted and the conclusion denied in this sense.

What should *count* as denial, and how is it to be related to assertion? At the very least, the denial of a propositional content together with its assertion must count as defective in this salient sense, since the argument from A to A is valid (so we have the IDENTITY condition for this sense of 'defectiveness'). WEAKENING is straightforward too, since if a position is not defective, any position with fewer assertions or denials is also not defective in that sense. For STRENGTHENING, note that it is a condition showing us when adding a denial *is* coherent. If $[\Gamma : \Delta]$ is coherent and we cannot coherently assert A, then we must be able to coherently deny it. We might say then that A has been *implicitly* denied in a position in which Γ has been asserted and Δ has been denied.

Taking the necessity and sufficiency of the coherence and logical consequence, we have the following connection between coherence and consequence: the claim that $[\Gamma : \Delta]$ is incoherent can be recast *positively* as the endorsement of the *sequent* $\Gamma \Rightarrow \Delta$. We will henceforth use this more familiar notation, but keep in mind that the validity of the sequent $\Gamma \Rightarrow \Delta$ is to be thought of as the verdict that asserting Γ and denying Δ is incoherent. With this formulation, the norms of coherence take a more familiar form as the structural rules of the sequent calculus.

IDENTITY: $A \Rightarrow A$

WEAKENING: $\dfrac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta} \quad \dfrac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow A, \Delta}$

STREGTHENING: $\dfrac{\Gamma \Rightarrow A, \Delta \quad \Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}$

So, the STRENGTHENING rule is a formulation of the usual rule CUT from the sequent calculus. From bottom to top, we strengthen positions by adding a statement, either to the left or the right. From top to bottom, we cut out that statement.

[2] Let us not beg the question in favour of multiple conclusion arguments at this point.

We choose coherence as a starting point, because it enables us to *do* logic in such a way as to result in classical logic (as we will see soon), it coheres with mathematical practice, and it does not require truth conditional or model theoretic semantics while still managing to be recognisably *semantics*. With this perspective, we can 'do logic' as soon as we have a discourse that is recognisably bound by norms of coherence. Mathematical discourse is clearly such a discourse. So, let us now consider how to use the notion of coherence to clarify *semantics*. We start with the connectives of propositional logic.

14.2.2 Connectives

Now consider the sequent rules for the propositional connectives. Here are the rules for negation and conjunction.

$$\textit{Negation}: \frac{\Gamma \Rightarrow A, \Delta}{\Gamma, \neg A \Rightarrow \Delta} \qquad \frac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \neg A, \Delta}$$

$$\textit{Conjunction}: \frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow A, \Delta \quad \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \wedge B, \Delta}$$

These clauses can be seen as ways to add vocabulary (here, propositional connectives) to a discourse, and to continue constraining that with respect to coherence. The negation rules tell you that if denying A is incoherent (in the context of asserting Γ and denying Δ) then so is asserting $\neg A$, and that if asserting A is incoherent (in the context of asserting Γ and denying Δ) then so is denying $\neg A$. Or to put it positively, if the assertion of $\neg A$ is coherent (along with asserting Γ and denying Δ) so is the denial of A, and if the denial of $\neg A$ is coherent (along with asserting Γ and denying Δ) so is the assertion of A. The rules for negation tell you how to treat the assertion of a negation and the denial of a negation, with respect to coherence.

Similarly the rules for conjunction dictate the behaviour of that connective with respect to coherence. If the assertion of $A \wedge B$ is coherent (in some context) then so is the separate assertion of A and of B. On the other hand if the denial of $A \wedge B$ is coherent (in some context) then either the denial of A is coherent (in that context) or the denial of B is coherent (in that context).

In this way, the rules for the connective tell you what to *do* with them. The traditional natural deduction rules for the connectives (infer $A \wedge B$ from A, B, and infer both A and B from $A \wedge B$) are instances of these rules: By identity, we have $A \Rightarrow A$ and $B \Rightarrow B$. It follows that $A, B \Rightarrow A \wedge B$. Similarly, $A, B \Rightarrow A$ (by weakening the identity $A \Rightarrow A$) so it follows that $A \wedge B \Rightarrow A$. Similarly, we have $A \wedge B \Rightarrow B$.

The result is classical propositional logic. (For example, we have $\neg\neg A \Rightarrow A$, via $\Rightarrow A, \neg A$.) As a matter of fact, I take it that there is a defensible natural deduction system in which *proofs* allow multiple premises and multiple conclusions [23], but discussing this would take us to far away from the present topic.

These rules have the attractive virtue that if we add them alone, then the relation of coherence defined on the new vocabulary satisfies the conditions of identity, weakening and strengthening. It follows that the addition of these rules is *conservative* over a base vocabulary without these logical connectives. (However, it is not necessarily conservative over a base vocabulary already containing logical connectives, such as a non-distribuitive pair of lattice connectives, or an intuitionistic conditional. There is significant debate over what this might mean [6].)

These rules tell us *something* about how to use the connectives. They do not tell you everything of what '$\neg$' or '$\wedge$' might 'mean,' but they do tell you how to use these connectives when it comes to evaluating assertions and denials featuring them, for coherence. In other words, rules such as these give you a starting point for the *practice* or *precisification* of a concept. For example, with these rules at hand we can see that a dialetheist or an intuitionist or a supervaluationist is not using negation this way when they diverge from these rules. A dialetheist may take $[p, \mathsf{not}\ p\ :\ ]$ to be coherent, and an intuitionist or a supervaluationist may take $[\ :\ p, \mathsf{not}\ p]$ to be coherent. It follows that they are not using their concept 'not' in a way that conforms to the rules above. However, it is another thing entirely to say that in the mouth of an intuitionist or a dialetheist or a supervaluationist, 'not' does not mean *not*. The sequent calculus rules are a very useful technique for constraining use, and for making precise a concept (in just the same way as we might present a truth table and say that we will take disjunction to behave like *this*). Sequent rules facilitate the introduction of connectives by a kind of *definition*. However, we well know that definition is not all that there is to say about meaning, as a definition might introduce a term into the vocabulary, and the vicissitudes of use might sweep it in another direction.[3]

This account of coherence and logical consequence does not appeal to *truth* or to *warrant*. It is semantically anti-realist in Dummett's sense [6], in that it does not take the preservation of warrant-transcendent truth—or indeed, any kind of truth—to be constitutitive of logical consequence. The approach is normatively inferentialist [3, 4] as the central notion (coherence) is evaluative, and it is understood in terms of the category of inference (at least, if we are prepared to understand the inference $A \Rightarrow B$ in terms of the incoherence of the assertion of A and the denial of B) rather than the category of representation.[4] This interpretation differs from the usual proof-first interpretation of intuitionist propositional logic: after all, the result is a defence of *classical* logic, and not *intuitionist* logic. The result is not the traditional BHK interpretation of intuitionist logic, in which the semantics of propositions is defined in terms of proof: a proof of a conjunction is a pair of proofs, one for each conjunct. A proof for a conditional is a function taking proofs of the antecedent to proofs of the consequent, and so on. This interpretation is well suited to the interpretation of logical consequence as preservation of *warrant*, and a conception of proofs according to which they have a number of premises and a single

[3] Consider the changes in the meaning of the terms *force* and *mass* in physics [11].

[4] For other work on the sequent calculus, assertion and denial: [12, 14, 19, 24].

conclusion. If we take proofs to have a more rich structure (allowing for multiple conclusions, as we have seen), or if we allow for richer operations on proofs (such as *continuations* [28], for one example), then the clauses for the connectives in the BHK interpretation motivate classical logic, rather than intuitionist logic.

14.2.3 Names, Variables and Quantifiers

If we wish to give an account of the logical features of the first-order quantifiers (and surely we must if we are to do justice to the logic of mathematics), then we go beyond the combination of propositions with other propositions and we analyse some of the internal structure of propositions. In the language of first-order logic, we compose propositions out of *predicates* and *terms*.[5] Some terms are *variables*, which play a role in quantified expressions. Some of the terms may play a special role in mathematical theories, such as the term '0' in arithmetic, or terms built up using function symbols, such as '$x' + (y \times z)$'. We wish to understand the addition of quantificational vocabulary in terms of the rules for the coherence of packages involving quantifiers. For the rules of the quantifiers to work, we need one last piece of terminology. We need *arbitrary* names, which are able to stand in assertions and denials. One way to think of arbitrary names is as *unbound* variables.[6] Think of the fragment of discourse: 'suppose x is a number, then x is either even or odd. If x is not odd, then ...' — 'x' here can be understood as an arbitrary name. If we wish to treat the fragment 'if x is not odd, then...' as a statement, with its own inferential properties—rather than as a component of a larger expression in which a quantifier binds the *variable* x—then an understanding like this seems appropriate. If you hold to an understanding of variables according to which they are always bound, you must choose a stock of names that have no inferential capacities of their own. What is means is simple: if a is an arbitrary name then whenever $[\Gamma{:}\Delta]$ is coherent, it would remain coherent with the replacement of some term occuring in Γ and Δ by the name a. (So, in the traditional vocabulary of arithmetic, 0 is not an arbitrary name but variables such as x and y behave as arbitrary names. In classical first-order *logic*, or higher-order logics, all names are arbitrary.)

Clearly, in mathematical practice, we have arbitrary names. We use them all of the time in reasoning when we make suppositions and reason under hypotheses. With this concept in mind, we can now examine the rules governing the universal quantifier.

$$\textit{Universal Quantifier}{:}\ \frac{\Gamma, Bt \Rightarrow \Delta}{\Gamma, (\forall x)Bx \Rightarrow \Delta} \quad \frac{\Gamma \Rightarrow Ba, \Delta}{\Gamma \Rightarrow (\forall x)Bx, \Delta}.$$

5 In what follows, we will use the following shorthand notation. For a formula B with some occurrences of a term t marked out, we will write 'Bt', and we will write 'Bs' for the result of replacing the selected instances of t in Bt by s.

6 Or, as Schütte calls them, *free object variables* [25].

In the second rule, we require a to be *arbitrary* and not present in $\Gamma \Rightarrow (\forall x)Bx, \Delta$. The motivation for these rules is clear. If the assertion of $(\forall x)Bx$ is coherent, then so is the asseriton of Bt, which is a specific case of the general claim. Conversely, if the denial of Ba is incoherent (where we have assumed *nothing* about the object a) then it is incoherent to deny that *everything* has property B.[7]

These rules give the usual properties of the universal quantifier. We have $(\forall x)Bx \Rightarrow Ba$ by the assertion rule and the identity $Ba \Rightarrow Ba$. By the denial rule, if we derive Fa for an arbitrary name a, then this derivation applies generally. Is there any *object* which we cannot therefore show to be B? Suppose there were: call such an object c. Run through the proof for Ba, except with 'c' in place of 'a.' We can do that since the name 'a' is arbitrary.

You may be concerned that the possiblity of non-denoting terms invalidate this account. Perhaps we should modify our rules by adding an 'existence' predicate as seen in free logics. Then we would say that $[\Gamma, (\forall x)Fx : \Delta]$ is incohrent if $[\Gamma, Fa, E!a : \Delta]$ is incoherent, and $[\Gamma : (\forall x)Fx, \Delta]$ is incoherent if $[\Gamma, E!a : Fa, \Delta]$ is incoherent for an arbitrary name a not in $[\Gamma : \Delta]$. I will ignore this complication, since non-denoting terms do not seem to play a significant role in mathematical discourse.[8]

This kind of account of the quantifier sidesteps the usual debate between 'objectual' and 'substitutional' accounts of the quantifiers. We have not relied upon an 'intended domain' of quantification, yet neither have we given a substitutional account identifying the *truth* of all $(\forall x)Bx$ with the *truth* instances of Bx. We have explained the inferential capacity of $(\forall x)Bx$ in terms of substitutions, not its *truth* [15].

You can do the same thing for higher-order quantifiers, using rules of the same general shape as the quantifier rules we have seen. We can also add very natural rules for identity [20], using arbitrary *predicates*, instead of arbitrary names.[9]

The quantifier rules have the usual harmony properties. For IDENTITY, we have $(\forall x)Bx \Rightarrow Ba$ (if we have at least one name). Since a is arbitrary, we have $(\forall x)Bx \Rightarrow (\forall x)Bx$. For STRENGTHENING, suppose $[\Gamma : (\forall x)Bx, \Delta]$ and $[\Gamma, (\forall x)Bx : \Delta]$ are both incoherent. This means that $[\Gamma, Bt : \Delta]$ is incoherent for some t, and $[\Gamma : Ba, \Delta]$ is incoherent for an arbitrary a. Then, $[\Gamma : Bt, \Delta]$ is incoherent (a has no distinctive coherence properties of its own) and hence, by strengthening for Bt, $[\Gamma : \Delta]$ is coherent.

---

[7] We require the restriction to *arbitrary* names a, since proving that 0 is even (on the basis of no assumptions) should not be enough to prove that every number is even.

[8] However, consider $\frac{1}{0}$ or $\lim_{x\to 0}\frac{1}{x}$. If these are not eliminated from the vocabulary, the something like Beeson and Feferman's logic of 'definedness' seems appropriate [7].

[9] In sequent form, using X as an arbitrary predicate, the rules are $\dfrac{\Gamma, Xs \Rightarrow Xt, \Delta}{\Gamma \Rightarrow s = t, \Delta}$ and $\dfrac{\Gamma \Rightarrow Bs, \Delta}{\Gamma, s = t \Rightarrow Bt, \Delta}$.

14.3 Mathematical Practice and Mathematical Theories

Now, let us turn to mathematical reasoning. Clearly, mathematicians *assert* and *deny*. Mathematical discourse is well suited (perhaps it is *uniquely* suited) to evaluation for coherence by the canon of classical first-order logic. Consider someone who engages in the practice of reasoning in the language of arithmetic. Let us be completely *agnostic* on the ontology of numbers. However, let us also take it that the discourse has the kind of form that it appears to take, on face value. A statement of the form '$2 + 3 = 5$' contains three terms, one function symbol and the relation symbol of identity. It is not a disguised statement about numerical quantifiers, and in the mouth of a mathematical reasoner it is asserted, and not merely 'play'-asserted. Accounts of the structure of mathematical statements that do not take them at something *like* their face value must explain what kind of structure they have, and how this structure suffices for the correctness of classical logic when reasoning with these statements.[10] So, we will take the analysis of the structure of mathematical claims at face value.

14.3.1 Introducing Mathematical Vocabulary

Here is how you can 'do arithmetic.' You can join in to the discourse of arithmetic by adding the term '0', the function symbol '$'$' and the predicate N to your vocabulary, and by following the following norms.

$$\frac{}{\Rightarrow N0} \quad \frac{}{\Rightarrow 0 = 0} \quad \frac{}{x' = 0 \Rightarrow}$$

$$\frac{\Gamma \Rightarrow Nx, \Delta}{\Gamma \Rightarrow Nx', \Delta} \quad \frac{\Gamma \Rightarrow x = y, \Delta}{\Gamma \Rightarrow x' = y', \Delta} \quad \frac{\Gamma, x = y \Rightarrow \Delta}{\Gamma, x' = y' \Rightarrow \Delta}$$

$$\frac{}{Nx \Rightarrow x + 0 = x} \quad \frac{}{Nx, Ny \Rightarrow x + y' = (x + y)'}$$

$$\frac{}{Nx \Rightarrow x \times 0 = 0} \quad \frac{}{Nx, Ny \Rightarrow x \times y' = x \times y + x}$$

[10] For example, consider an approach that uncovers the 'meaning' of a mathematical statement in terms of a conditional (if there is an ω sequence, then ...). You must show, for example, that the statement (if there is an ω sequence then $\neg A$) should either be equivalent to the negation $\neg$(if there is an ω sequence then A) or there should be an explanation of the divergence, for the mathematical statement $\neg A$ *appears* to be the negation of A, but on the conditional analysis of mathematical statements, appearances are deceiving.

$$\frac{\Gamma \Rightarrow B(0), \Delta \quad \Gamma, Nx, B(x) \Rightarrow B(x'), \Delta}{\Gamma, Nx \Rightarrow B(x), \Delta}$$

This is a way to *add* arithmetic vocabulary to your already existing inferential repertoire. (This is why we included the predicate N for 'is a number'.) What we have is a sequent system for Peano Arithmetic, with one exception. We have placed no restriction on the judgement $B(x)$ to appear in the induction rule. $B(x)$ can be *any* statement, as long as the x is arbitrary.

What is the upshot of treating the new vocabulary as constrainted by these rules of coherence? The first feature making this differ from any *logic* you have already seen is that by following these rules, you are committed to $\Rightarrow (\exists x)(x = 0)$. We have *proved* that there is a number. What else can you prove? If your vocabulary merely contains arithmetical terms, then you will be committed to Peano Arithmetic (PA). If the predicates used in induction rule range more widely, you may commit yourself to something stronger.

Can we govern coherence in this way? Is it legitimate to 'define' objects (like the number zero) into existence? It makes sense if we recall the treatment of other 'definitions'. Recall the connectives. If we find someone who is prepared to assert A, B but to deny $A \wedge B$, then we may be confident that the person is using '$\wedge$' in a non-standard way. The same goes with the predicate N and the term 0. If someone rejects $N0$ (as opposed to *eschewing* the vocabulary in which the claim is couched), then we may hold that she either has made a mistake, or she has not understood the predicate 'N' or the term '0'.

Now consider the benefits of understanding mathematical theorising and practice in this manner. It is telling you something about the *significance* of mathematical discourse, without explaining this using some particular model or class of models. The idea is not ontological economy[11] it is the direction of explanation. I shall discuss this further, below.

So, let us suppose that I teach you how to use arithmetical vocabulary according to these rules. You now judge arithmetical discourse using these canons of reasoning. (You take the denial of $2+2 = 4$ to be incoherent. You are committed to the validity of induction over all predicates in the vocabulary, etc.) Is there anything *else* needed for you to become a competent user of arithmetical vocabulary?

We can put the point colourfully: suppose we have two users of arithmetical vocabulary, equally competent with the rules we have considered, one of whom is in touch with '*the* ω sequence' and the other who is not. How would this difference manifest itself? How could we tell that we are reasoning like the one or the other?

Once you have added numerical vocabulary, there is no reason to stop there. You could add new prinicples to treat analysis, second order quantification, theories of

---

[11] Think about it: the formal language itself provides us with an *omega* sequence of formulas. We already have enough ontology when we have a language to speak. Numbers add no more.

sets. I will not go into detail on this, but the aim ought to be to find a natural smooth axiomatisation of the extended mathematical theory, that goes as far as possible towards fixing the behaviour of the new vocabulary.

14.3.2 Consistency and Conservativity

This addition of arithmetical vocabulary is not necessarily conservative over your pre-arithmetical vocabulary. As an extreme case, consider a finitist, committed to $(\forall x)(x = a_1 \vee \cdots \vee x = a_n)$ for some n. Given that you can prove the existence of more than n numbers, the theory becomes inconsistent with the addition of the rules. However, it is possible to regain conservativity, if you are a little more careful in the way in which you add your language, even this is possible. Divide the terms into two *types*. We have two forms of quantification, one for each type. Add the numerical vocabulary to the language as inhabitants of a new type, so you do not substitute an arithmetical name in the positions taken by names in the original discourse, and you don't substitute your original names in your arithmetical positions. You have two kinds of quantifiers, the quantifiers of arithmetic, and those of the original discourse. If arithmetic is consistent, then this addition is conservative: adjoin to any model of the old language, a disjoint domain of numbers, and let one set of quantifier range over the old domain, and the other over the new domain of numbers. This is a model of the expanded language, and the interpretation of the formulas from the original language is unchanged. So, the addition of numerical vocabulary and arithmetical theory in this way is conservative, if consistent.

Of course, it is possible to add a truly expansive *existential* existential quantifier, that binds an *ambiguous* variable which may be substituted into a spot appropriate for a variable of either type. The question then arises: which of the existential quantifiers is the appropriate one? Why choose a bifurcated language over one with a single categoy of object variables and a single category of objectual quantifiers?

Here is the view concerning the behaviour of mathematical theories and their semantics. We adopt a mathematical theory by (1) introducing new vocabulary (2) constraining our patterns of assertion and denial in that vocabulary in such a way that it (3) remains conservative over the pre-mathematical vocabulary.

14.4 Consequences of the View

Now I'll chart out consequences of the view, by comparing it to a number of extant positions.

14.4.1 Realism and 'Platonism'

On this view, to *use* arithmetical vocabulary is to commit yourself to mathematical ontology. Given that you use the vocabulary in this way, there is no sense in which you need further information as to whether or not there are any numbers. They exist, mind independently and necessarily. It is a kind of 'thin realism' to use Maddy's vocabulary [18], combined with a semantic anti-realism. Given that existence is what is expressed by the existential quantifier, this is the natural and default position if we take the grammar of mathematical claims seriously.

(We have nothing against the project of people who tell us that there is a *stronger* notion of existence, which is not shared by mathematical objects. It remains for the proponent of these positions to articulate the kind of *stronger* notion of existence that they have in mind.)

The semantic anti-realism of this position means that we may use the machinery of the classical sequent calculus without starting with a notion of a *model* in which arithmetic must be interpreted, or by presuming that there is a structure out of which arithmetical claims inherit their truth. As a matter of fact, there *is* a model of our arithmetical theory (at least, you will see that there is, once you adopt the vocabulary of sets), but to say that arithmetical claims are true *because* of their relationship to some particular structure is to get the order of explanation backwards.

We agree with Platonism about the existence—and even, perhaps the *necessary* existence, and mind-independence, etc.—of mathematical objects. We part from Platonism concerning the direction of explanation of mathematical truth. One need not explain the significance of mathematical vocabulary by way of reference or representation of mathematical objects. To explain the truth of a mathematical claim in terms of the properties of the mathematical objects referred to in that claim is to not offer the only kind of explanation ... (and in fact, to leave the core concept unexplained at all). Rather, another explanation is possible, in terms of coherence of the assertion and denial of mathematical claims: in other words, we take the *proof* to be the explanation.

14.4.2 Formalism

It might seem to follow that since we have identified mathematical explanation with *proof* that I am committed to a kind of formalism, since a proof here is a formal proof in a deductive system. And formalism, at least as far as the commitment to formally articulated proof as the means for establishing mathematical truth is concerned, is plausible.

However, the view articulated here does not mean that arithmetic is merely a particular first-order theory such as Peano Arithmetic (PA). One can prove facts about arithmetic using non-arithmetic vocabulary, as predicates in non-arithmetic vocabulary may be substituted into the induction rule as necessary.

So, for example, if arithmetic is embedded into a reasonably strong set theory (ZF will do), then, substituting sentences involving set membership and other predicates not definable in arithmetical vocabulary into the induction scheme will enable us to prove more claims—in arithmetical vocabulary—than we can otherwise. As one example, we can prove CON(PA) (which can be expressed in a first-order sentence in the language of PA) by using stronger induction schemes allowing for induction over sets as well as first-order sentences in arithmetic vocabulary.

Similarly, people have objected to logicism on the grounds that mathematical theorems are typically not proved inside an easily isolated formal system. Wiles' proof of Fermat's last theorem, uses a great deal of abstract mathematics beyond the first order theory of arithmetic. It seems that we cannot specify the formal theory in which Wiles proof obtains [27]. The response is straightforward: the rules of arithmetic *explicitly* give you a place to import other vocabulary. Arithmetic as I have defined it is open ended, depending on what *else* is in our vocabulary. If we extend our language to contain a new predicate B, and we did *not* concede that induction with respect to B worked, then this would be akin to not adding instances of *modus ponens* with respect to our new vocabulary when we expand our language.

The fact that the induction schema contains a space for *arbitrary* induction predicates means that our arithmetic theory is not static and fixed: it grows as our language grows. It cannot be identified with a particular first-order theory such as Peano arithmetic.

This does not mean that PA does not play a special role. Dan Isaacson holds that PA delineates the class of genuinely *arithmetic* truths [13]. For us, the reasoner who is committed to the vocabulary of arithmetic (and nothing else, or at least, nothing else that she can import into the induction rule) can prove only what is provable in PA. By her lights, denying CON(PA) is coherent, and so is asserting PROV$(0 = 1)$, at least when these are construed as sentences in the language of arithmetic.

14.4.3 Hilbert

This view is recognisably in the tradition of David Hilbert, because there is an important sense in which consistency (together with formality, which is required for conservative extension) is all that is required for mathematical existence. We may do for other mathematical theories what we did for arithmetic. If a mathematical theory (of sets, of categories, of whatever else) is consistent, we may add the new vocabulary to our own, giving rise to a richer vocabulary, conservative (if we are careful) over the old theory. This much of Hilbert's program is worth keeping.

Of course, Hilbert's program of finding certainty through finitist consistency proofs is dead. There are no finitist consistency proofs for any interestingly strong mathematical theory. This does not mean nothing to the general Hilbertian insight that there is no more to mathematical existence than consistency.

14.4.4 Carnap

This view is *Carnapian*, since the perspective of mathematical theories allows us to distinguish the internal and the external questions concerning mathematical existence. The question concerning whether there *really are* any numbers is answered *internally* by the user of the vocabulary (who is genuinely *asserting* and *predicating* and the like) in the affirmative. She can prove that $(\exists x)(x + 5 = 12)$, so 7 *exists*. (Unless, of course, you have a stronger reading of existence claims, according to which we can existentially quantify over non-existent objects.) The *external* question is a different matter, and like Carnap, we answer the external question of whether or not we ought adopt a mathematical theory on pragmatic grounds. The nature of your answer will depend on the precise version of the question asked.

It's only a *modest* Carnapianism, and it is not refuted by Gödelian worries which spelled the end of Carnap's own program [10, chaps. 7 and 9]. For Carnap, there was taken to be a theory-independent and neutral analytic–synthetic distinction, a neutral perspective from which you could judge what was analytic in a theory. As we can see, the question of what is provable in a theory like PA requires *more* than PA to articulate, not less. Metamathematics is more mathematics, not a retreat from mathematical commitment.

Is this view beholden to a pernicious or implausible version of the analytic–synthetic distinction? It does not seem so. We do not need to identify the meaning of an expression with the rules governing coherence of assertions and denials involving the expression. We merely need to say that we can *introduce* (or *explicate*) vocabulary by treating it as constrained by some collection of rules for coherence.

14.4.5 Plenitudinous Platonism, and Fictionalism

Similarly, the more recent view of 'plenitudinous platonism' holds that there the mathematical universe is as full as it can be. It is quite difficult to characterise 'plenitudinous platonism' [2, 21]. The motivating picture is that any kind of mathematical structure that *can* exist *does* exist. Our position provides a plausible reconstruction of the idea: any consistent mathematical theory may (if we like) be *adopted*, enriching our own mathematical vocabulary. The universe places no limit to the extent of mathematical theorising.

As Balaguer notes [2], fictionalism about mathematics is structurally quite similar to plenitudious platonism. For fictionalism, mathematical theories are made up, and we never need go to check that there are objects that the theories are talking about—beyond assuring ourselves that the theory is consistent [8, 9]. So far, we agree with fictionalism. However, instead of taking the posited theory to be *fictional*, we can take them to be true. There is no need to take mathematicians to be *mistaken* except for an overactive sense of ontological economy.

14.4.6 Structuralism

This view is *structuralist* [26] because the only general assurance that a mathematical theory is a conservative addition (if consistent) is when the new vocabulary is completely disjoint from our old vocabulary. A mathematical theory cannot be *about* cows or tables or chairs or whatever else we are talking about when we are not doing mathematics. It may, on the other hand, be *applied* to such things, by taking deductions and conclusions couched in mathematical vocabulary and applying them elsewhere. (Arithmetical facts are applied when we count cows or pay the bills, topological facts may be used in discussions of the large-scale structure of the universe, and there are many other applications besides.) Mathematical facts (the kinds of things to which we are committed in *using* mathematical vocabulary) are structural because they are reapplicable. In one sense, any *arithmetical* claim, because it is founded on a very simple base (i.e. the rules of arithmetic) may be re-applied to any structure on which those rules may be reinterpreted. In this sense, we have structuralism without having to give an account of what a mathematical structure is, since we have an alternative explanation of the meaning of mathematical vocabulary.

14.5 Miscellaneous Concluding Remarks

14.5.1 Using and Mentioning

The distinction between *adopting* a mathematical theory and *exploring* a mathematical theory plays an important role. We may not want to adopt *all* mathematical theories as they come up, in the same kind of way as I urged you to adopt arithmetic. We can, for example, adopt something rather strong such as ZFC and then *interpret* claims about, say non-wellfounded sets as claims about *graphs*, which themselves are thought of as particular sets in ZFC (ordered pairs consisting of a carrier set and a relation on that set) [1]. You learn ZFA (ZF with the *anti*-foundation axiom) by translating it into your native tongue. You *can* do this, but you may find that when you do so, you begin to speak ZFA like a native, and cease to translate it.

On the other hand, there is much to be said for keeping your mathematical vocabulary (in essence) very small by adopting a set theory such as ZFC and translating other vocabulary into it as necessary. This ensures that the addition of new vocabulary (if interpretable within your set theory) will not come at the cost of consistency given that you have already paid the price of adopting your favoured theory of sets.

Similarly, you can do mathematics of particular structures without adopting the vocabulary of that discourse at all. You can do it by *mentioning* the vocabulary and not *using* it. You could (using the language of *formulas* and *proofs*) consider whether or not a particular sentence follows from some set of axioms. You could, if you wish, say that '$2+2=4$ is a theorem of PA' without taking it that $2+2=4$. (This may be the attitude of mathematicians exploring set theoretical axioms that they do not take to be plausible [16, 17].)

14.5.2 Ontology and Epistemology

It answers the ontological question of the existence of mathematical objects in two ways. Firstly, given the vocabulary that we *use*, the internal question has a straightforward answer. There are numbers. There are sets. They exist necessarily and independently of us. (That last claim is not a part of the mathematical theory. It will follow from a decent theory of modality and dependence.) However, we do not need to explain mathematical knowledge by means of 'contact' with the realm of mathematical objects. The *general* question (what about new kinds of mathematical objects that we haven't considered?) can only be answered piecemeal. It seems that whatever language we adopt, we can add *more*. (Reflection principles seem to ensure that whatever mathematical theory we adopt, it may be extended. CON(T) adds new sets/structures over T.)

The epistemic question of how we come to *know* mathematical truths also has a two-track answer. Given particular mathematical concepts we may draw consequences on the basis of traditional deductive argument. The more interesting question is why we use concepts such as the ones that we have in the ways that we do. For this, different kinds of answers are available. A *pragmatic* answer will explain the choice of some vocabulary rather than another. This seems to do justice to the kinds of discussions set theorists have concerning open questions such as the continuum hypothesis.

Consider the position of the mathematician exploring the theory of sets. The best theory commits us to GCH $\vee$ $\neg$GCH, but it seems that it leaves open which disjunct is true. Contemporary set theory is a complicated affair in which the search is on for different considerations in favour of GCH or $\neg$GCH. The set theorist is attempting, in these circumstances, to articulate and sharpen our account of the concept 'set' in ways that satisfy sensible desiderata, such as the goal to MAXIMISE the set-theoretical universe [16–18]. This is quite sensible, given the set-theoretical goal of finding a large home (or 'vocabulary') in which to interpret or translate all different kinds of mathematics. The kind of *freedom* involved in this exercise explains both the appeal and the coherence of staggeringly large cardinal axioms.

Acknowledgements This paper has a webpage: http://consequently.org/writing/antirealist. Check there to post comments and to read comments left by others. Thanks to Dan Isaacson for hospitality during my visit to Wolfson College and the Philosophy Faculty at Oxford University, and to Allen Hazen, Dan Isaacson, Øystein Linnebo and Alexander Paseau, the audience at the (Anti)Realism conference in Nancy, and two anonymous referees, for helpful comments and discussion. This research is supported by the Australian Research Council, through grant DP0343388.

References

1. Aczel, P. 1988. *Non-Well-Founded Sets*. Number 14 in CSLI Lecture Notes. Stanford, CA: CSLI Publications.
2. Balaguer, M. 1998. *Platonism and Anti-Platonism in Mathematics*. Oxford: Oxford University Press.
3. Brandom, R. B. 1994. *Making It Explicit*. Cambridge, MA: Harvard University Press.
4. Brandom, R. B. 2000. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.
5. Coffa, J. A. 1993. *The Semantic Tradition from Kant to Carnap*. Cambridge, MA: Cambridge University Press. Edited by Linda Wessels.

6. Dummett, M. 1991. *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.
7. Feferman, S. 1995. "Definedness." *Erkenntnis* 43(3):295–320, 11.
8. Field, H. 1980. *Science Without Numbers: A Defence of Nominalism*. Oxford : Blackwell.
9. Field, H. 1991. *Realism, Mathematics and Modality*. Oxford: Blackwell.
10. Friedman, M. 1999. *Reconsidering Logcal Positivism*. Cambridge, MA: Cambridge University Press.
11. Friedman, M. 2001. *Dynamics of Reason: The 1999 Kant Lectures at Stanford University*. Stanford, CA: CSLI Publications.
12. Hacking, I. 1979. "What Is Logic?" *The Journal of Philosophy* 76:285–319.
13. Isaacson, D. 1992. "Some Considerations on Arithmetical Truth and the ω-Rule." In *Proof, Logic and Formalization*, edited by M. Detlefsen, 94–138. London: Routledge.
14. Kremer, M. 1988. "Logic and Meaning: The Philosophical Significance of the Sequent Calculus." *Mind* 97:50–72.
15. Lance, M. 1996. "Quantification, Substitution, and Conceptual Content." *Noûs* 30(4): 481–507.
16. Maddy, P. 1988a. "Believing the Axioms 1." *Journal of Symbolic Logic* 53:481–511.
17. Maddy, P. 1988b. "Believing the Axioms 2." *Journal of Symbolic Logic* 53:736–64.
18. Maddy, P. 2005. "Mathematical Existence." *The Bulletin of Symbolic Logic* 11(3):351–76.
19. Price, H. 1990. "Why 'Not'?" *Mind* 99(394):222–38.
20. Read, S. 2004. "Identity and Harmony." *Analysis* 64(2):113–15.
21. Restall, G. 2003. "Just What *Is* Full-Blooded Platonism?" *Philosophia Mathematica* 11: 82–91.
22. Restall, G. 2005. "Multiple Conclusions." In *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, edited by Petr Hájek, Luis Valdés-Villanueva, and Dag Westerståhl, 189–205. London: KCL Publications.
23. Restall, G. 2007. "Proofnets for S5: Sequents and Circuits for Modal Logic." In *Logic Colloquium 2005*, Number 28 in Lecture Notes in Logic, edited by Costas Dimitracopoulos, Ludomir Newelski, and Dag Normann. Cambridge, MA: Cambridge University Press.
24. Rumfitt, I. 2000. " 'Yes' and 'No'." *Mind* 109(436):781–823.
25. Schütte, K. 1977. *Proof Theory*. Berlin: Springer. Translated from the German by J. N. Crossley.
26. Shapiro, S. 1997. *Philosophy of Mathematics: Structure and Ontology*. Oxford: Oxford University Press.
27. Shapiro, S. 2000. *Thinking About Mathematics: An Introduction to the Philosophy of Mathematics*. Oxford: Oxford University Press.
28. Streicher, T., and B. Reus. 1998. "Classical Logic, Continuation Semantics and Abstract Machines." *Journal of Functional Programming* 8(6):543–72.

Chapter 15
A Tale of Two Anti-realisms

Sanford Shieh

15.1 Epistemological Anti-realism

I'm going to start by recounting a familiar story about anti-realism and classical logic. The anti-realism that figures in this story is the kind inspired by Michael Dummett's writings. I should also note that I'm giving you an old-fashioned story, not one of those newfangled fictionalist fictions, so I'm going to omit a lot of details that would be needed to achieve the pretense of truth.

Anti-realism is founded on a doctrine about linguistic meaning. In order to communicate in language, we have to know what our conversational partners mean by the words and sentences that they utter. If someone says to me, "Your scarf is in the boot," she will not have succeeded in telling me what she wanted to tell me if I think she's referring to a kind of shoe when she says "boot" but in fact she means a part of a car. What she wanted to tell me is something about the world. My mistake about what she means by the word "boot" leads to a failure to know what she's telling me. This is because I wrongly think that the truth of the sentence she uttered depends on conditions obtaining in some shoe, and so I take her to have expressed a belief about those conditions, while, in fact, she meant to tell me about conditions in some car.

But how do I figure out what she meant? What sort of evidence is available to me for the meaning that she associates with that word? Or for the conditions that she intended to describe with that sentence? Surely I should start by finding out the ways in which she uses these words. In particular, I would try to figure out what sorts of things she would call a "boot", i.e. what sorts of things she would take to be correctly described by that word. And, I should try to figure out what sorts of circumstances she would take to be correctly described by her sentence.

How would we go about gathering such evidence? Let's consider how we might find out what someone means by the word "gasket". If in the presence of the entrails of a car he can point to something and say that it is a gasket, then we have some

S. Shieh (✉)
Department of Philosophy, Wesleyan University, Middletown, CT 06459, USA
e-mail: sanford.shieh@wesleyan.edu

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_15,

basis for thinking that what he's pointing to is what he means by "gasket". Only some basis, of course. Quine and Davidson, among others, have taught us just what an intricate business it is to infer meaning from observations of peoples' utterances and features of the environment in which they make these utterances. But, so the story goes in any case, evidence of this sort is ultimately all we've got to go on; it's all the hard data, as it were. So if we are to understand one another, it must be possible to observe our uses of at least some words and sentences to describe the world.

We're not quite at anti-realism yet. Let's go back to "gasket". What if our test subject can't actually tell gaskets from non-gaskets? What if there isn't anything that he can recognize as correctly described as a gasket? Maybe all he knows is that gaskets are parts of cars. But, pressed by a persistent philosopher, he might send her to talk to the guys at the *Jiffy-Lube* who work on his car. Well, what if they also can't tell gaskets from non-gaskets? (Maybe changing oil and filter no longer involves knowing anything about gaskets....) So perhaps then the philosopher would have to look it up in a dictionary, or on Google, or go to some automotive engineering department in some university, or a car manufacturer. But could all these attempts fail? What if no attempt of this sort succeeds? Is that possible? Could it be that nobody can tell gaskets from non-gaskets? Could it be that there are such things as gaskets, and people do use the word "gasket" to refer to them, but no one who uses the word can recognize anything as correctly called "gasket"? If that's the case, how can I ever find out what all these people mean by "gasket"? As already mentioned, all I've ultimately got to go on is how they use the word. So now it seems that although they utter the word alright, these utterances make no contact with the world. So it's hard to see what in the world they are talking about when they talk about gaskets.

The conclusion, then, is that if what someone means by a word is to be communicable, then whatever it is that the word correctly describes must be something that she, or someone in her linguistic community, can recognize as correctly described by that word. Their exercises of this recognitional capacity manifests, in the sense of furnishing evidence for, the meaning she associates with the word.

The same sort of thing holds of sentences. If the condition that someone depicts by uttering a sentence is to be communicable, i.e. if it is to be possible for another person to know what that condition is, then the speaker, or someone in his linguistic community, must be able to recognize that the condition obtains if it indeed does, and recognize that it doesn't obtain, if indeed it doesn't. A speaker's exercise of this recognitional capacity provides evidence for, that is, manifests, his taking the sentence to describe these conditions. We have arrived at the fundamental thesis of anti-realism: communicable truth conditions must be manifestable.

Why is this anti-realism? The fundamental thesis stands in some tension with an intuitive idea that underlies metaphysical realism: reality is completely independent of our knowledge of it; the way the world is, the circumstances that actually obtain, might be completely unrecognizable by us. But such circumstances, according to anti-realism, cannot be conditions that we take our statements to express—in any event if they were no one would know that we expressed them.

Before we get to the classical logic part of the story, I should mention another version of the part of the tale I just recounted. There are those who would worry that the picture of language from which we had started is all wrong, because it presupposes[1] that the association of a mental content with linguistic expressions is central to the ability to speak a language. Having mastered some *Philosophical Investigations*, they think that speaking a language should be conceived, in the first place, as having a practical ability, much like the ability to touch-type. Most touch typists are not able to tell you the positions of the keys on keyboards on which they exercise their skills. They don't have what is called an explicit mastery of a set of propositions about these positions by knowledge of which they derive their typing. But their practical ability can be modeled or represented by these propositions, in the sense that someone who had explicit knowledge of those propositions can, if they think real fast and move their fingers equally fast in accordance with the conclusions they reach, duplicate the output of the touch-typist. That is, crediting a touch-typist with what is called implicit knowledge of propositions about keyboard layout constitutes a hypothesis by which one could explain their practical typing abilities. Such hypotheses are, however, constrained by Occam's Razor: one shouldn't ascribe any piece of implicit knowledge that is not required to account for some aspect of the subject practical abilities. The parallel to language, then, is that implicit grasp of the propositions of a theory of meaning is a theoretical posit by which we explain a speaker's capacities for verbal behavior. The picture, then, is that to know what you mean when you speak a language, I formulate fairly elaborate hypotheses about your unobservable states of knowledge in order to explain your (directly observable) linguistic (and other) behavior. But such hypotheses should not be multiplied beyond what's necessary to account for those behavioral capacities. So, again we have arrived at the fundamental thesis of anti-realism: implicitly grasped truth conditions must be manifestable in verbal behavioral capacities.

It should be clear why this is just another version of the same story. In each case an epistemic achievement is taken to be central to the phenomenon of language, and the idea of manifestation is explained in terms of how we can attain that achievement.[2]

15.2 The Bivalence Argument

Finally we get to classical logic. Here's a familiar and widely accepted explanation of what makes a form of reasoning deductively valid. Forms of argument are specified in terms of schematic statements abstracted from 'concrete' statements figuring in actual reasoning. The basis of the abstraction is the fact that in many ordinary statements we can discern parts whose specific meanings make no difference to the

---

[1] For more on presupposition theory see [2] and references therein.

[2] It's possible that some readers, at this point, would be filled with a sense that the foregoing just isn't the right account of the notions of implicit and explicit mastery of meaning. I invite them to say precisely why they think so. Alternatively, I invite those readers to take a look at [3].

way in which these statements are determined as true or as false. The results of the abstraction displays the forms of how "concrete" statements are determined as true or false according to their composition out of logical and non-logical expressions. Those aspects of the component expressions of a statement that are not abstracted away, i.e., that are sufficient for fixing the statement's truth or falsity, are their semantic values. The partial truth conditions of schemata are thus given by two things. The logical constants specify the functional dependence of the truth-value of a statement of that schematic form on the semantic values of its component expressions. Each possible set of semantic values for the component expressions then fixes a truth condition. Such a set of semantic values is a semantic interpretation. Semantic logical consequence, of course, holds from a set of schemata to a schema just in case no semantic interpretation verifies all the premise-schemata and falsifies the conclusion-schema.

We can separate out three stages or parts of this explanation of valid argument.

First, we have an uncontroversial specification of deductive validity. A form of argument is valid just in case it satisfies the following condition:

> If the truth conditions of any set of instances of the premises are fulfilled, then so is the truth condition of the corresponding instance of the conclusion. (V)

Second, on the basis of the meanings of the logical constants we have an account of the truth and falsity of schematic statements as a function of the semantic values of its constituent schematic letters.

Third, we have an account of all the possible semantic values of the non-logical expressions that can replace these schematic letters.

The last two stages together yield partial truth and falsity conditions of all statements of a schematic form, and, therefore, a further specification or interpretation of condition (V). One might call it the semantic value interpretation of (V). This interpretation is the explanation of valid argument.

Let's ponder very briefly in what sense we were asking for (and have now) an *explanation* of deductive validity. One idea is that we are seeking an analysis of the concept of valid argument. If so, then we must think that no argument could be valid unless it satisfies the terms of the explanation. Also, it seems to follow that in the explanation at hand, semantic values must be properties of expressions that they possess independently of the inferential relations of statements that contain them.

In this framework, classical forms of reasoning such as the law of excluded middle or the rule of double negation elimination are valid because of two factors: the meanings of the logical constants of negation and disjunction, and the claim that all sub-statements have as their semantic values exactly one of the two properties of truth and falsity.

The second factor is of course the principle of bivalence. In our present story, metaphysical realism justifies this principle. The idea is that if we have succeeded in expressing a fully determinate condition of the world with a statement, then that condition either in fact obtains, or it doesn't. So bivalence applies to any statement whose meaning isn't defective in some way.

How does anti-realism lead to a criticism of classical logic? According to anti-realism, if the truth conditions of a statement either obtain or fail to, then either we can recognize that they do, or we can recognize that they don't. So, in order to know that these conditions is in one of the two states, one would have to know that we have one of the two recognitional capacities. Hence if we don't know that we have one of the two recognitional capacities, we also don't know that either the conditions obtain or they don't. Which is to say, we don't know that the statement is either true or not true.

This leads us to the idea of an undecidable statement, a statement just beyond the limits of our known cognitive capacities, and so we don't know that we can come to recognize it as true, nor do we know that we can come to recognize it as false. (Clearly this is none of the notions of undecidable statement current in mathematical logic.) Anti-realism implies that if there are such statements, we don't know that they are either true or not true. The schema that is the law of excluded middle is true under all interpretation only if all unnegated disjuncts of its instances are either true or false. So, if some of these disjuncts are undecidable statements, we do not know that the schema is true under all interpretation, i.e., we don't know that it is logically valid. Note that this conclusion is, of course, not the claim that we know that the law of excluded middle is not logically valid. This latter claim is not anti-realism's ground for rejecting classical logic.

Let's call the this argument "the bivalence argument."

As I said, this is a familiar story. There are many intricate variations in the telling of it, but they are all based on three underlying themes. The first is that what motivates anti-realism is a problem in the epistemology of meaning: how do we attain knowledge of what someone means by her utterances? The second theme is that the solution to the epistemological problem imposes what we might call epistemic conditions on the nature of the objects of knowledge. I use the term 'epistemic condition' advisedly. This theme has a Kantian ring to it; it is an inference from the conditions for a certain kind of knowledge to be possible to properties of the objects of knowledge. The final theme is that anti-realism leads to a rejection of classical logic by undermining the claim that undecidable statements satisfy the principle of bivalence.

15.3 Conceptual Anti-realism

What I want to do now is tell a different story about anti-realism and classical logic. I'm not sure that it is a better story. Indeed sometimes I'm not altogether sure that it is a completely different story. But it is, I think, somewhat less familiar.

The basis of the type of anti-realism I am now sketching is not an epistemological issue but a conceptual one. Specifically, this anti-realism stems from an analysis of some connections among the concepts of meaning, of justification, and of assertion.

Let start by going back to the beginning of the last story: we use language to communicate. How do we do this? What is it for you to tell me something? It's a well-established point that this requires more than simply your uttering a declarative, grammatical sentence. For we may be in a play, or you might be telling me a joke, or a bedtime or barroom (but not a philosophical?) story. What's the difference? Why are the things that we say as part of those activities not assertions? One answer is that in those activities, in contrast to the activity of making assertions, we are not supposed to be aiming to say true things. Of course, just because we're supposed to be trying to say true things when we make assertions, it doesn't follow that whenever we make an assertion we do in fact intend to say something true. But lies uttered with the intention of saying something false wouldn't mislead unless their audiences took them to be aimed at the truth. If I'm telling you a joke I would be hard put to understand you if you were seriously to accuse me of lying to you. And, of course, people do try to wriggle out of lies by claiming to have been joking.

If assertions are supposed to be aiming at the truth, then in making an assertion I have to acknowledge in some way that I'm trying to say something true. How exactly do I make this acknowledgement? Well, if someone is really trying to say something correct, then it ought to matter to her whether what she says is indeed correct. So, it ought to matter to her whether there are reasons to think that what she says is correct. Just as, if I am trying to go to Brooklyn, then it ought to matter to me whether what I'm doing will take me to Brooklyn. This, I take it, is behind the thought that if someone can be taken to be making an assertion, then she ought to accept the legitimacy of requests to produce grounds for what she says. If she fails to accept this, then it is unclear that she takes reasons to think that what she says is correct to have anything to do with what she is saying. And then it's unclear that she acknowledges that she should be trying to say something correct. Similarly, if she is given reason to think what she says to be incorrect, but she doesn't accept this as a prima facie reason to "take back" what she says, it is again hard to see that she is concerned with the correctness of what she says.

So, in order to qualify as having made an assertion, a speaker must recognize the legitimacy of a request to produce grounds for the truth of the statement asserted, and be prepared to withdraw the assertion, should she be presented with considerations which she recognizes as showing that there are not sufficient grounds for taking the statement as true, or that there are actually grounds for taking the statement to be false. Another way of putting it is that, to be taken as making an assertion, a speaker must take what she says to be subject to assessment as correct or incorrect, by reference to what she *would count* as justifying it.

So far I've discussed assertion as a communicative practice. Next I want to say something about the connection between assertion and meaning. Let's begin with the intuitive distinction between genuine or substantive disagreements and merely verbal disagreements. Suppose that, in response to my friend's statements about my scarf I look in the boot I thought she was talking about and find it empty. I might say to her, "You're wrong, my scarf is *not* in the boot." Indeed, if I were trying to be obnoxious I might bring the shoe in question to her and point, in Sartrean fashion, to the absence of my scarf therein. If I were a really aggressive type of fellow I might add, "See, you should take back what you said." If we're still on speaking terms at that point, she might tell me I was looking at the wrong place, and open up the car to reveal my scarf to me. Should we really care to do so, we might discuss this little contretemps and come to see that we weren't really disagreeing. We could come to see that we were both right, and that neither of us has to take back what we asserted. And all this is so, because we didn't mean the same thing by the word "boot"; our disagreement was merely verbal. Now, if we weren't *really* disagreeing, why do we still characterize the episode as involving a verbal *disagreement*? The reason is that, at the outset, my friend took the assertion made with the sentence she used be correct, and I denied that the assertion I understood to have been made with that very same sentence is correct or justified. The disagreement lies in the fact that our attitudes to the assertion that each of us took to be made with a sentence are mutually contrary.

I now argue that it follows from the two things I've just discussed—the justification sensitivity of assertion and the role of meaning is genuine versus verbal disagreement—that we have difficulty imagining how two people could agree completely in what they would count as justifications for a statement but attach different meanings to it.

I'll do this by telling you a little story about two friends, Albertine and Odette. One day, Albertine tells Odette, "Charles now lives on Broadway," and Odette disagreed, or seemed to disagree. But, being friends, they talk about this (apparent) disagreement. In the course of the discussion, it turns out that Albertine's reason for saying what she said is that she had seen letters sent to Charles with a Broadway address, and had also seen, a number of times, Charles going into an apartment building on Broadway at night and leaving it on the following morning. Odette disagreed, on the other hand, because she heard from Charles himself that he had just moved to Amsterdam (the Avenue, not the city).

A true friend is another self; since Albertine and Odette are true friends, they agree completely in what they would count as reasons for and against saying, "Charles lives on Broadway." Still, we can imagine a number of different ways in which their chat about Charles might continue. For example, they might give more weight to what Charles himself says than to what they observe. If so, they might arrive at the plausible conclusion that Albertine had insufficient reasons for saying, "Charles lives on Broadway," while Odette's basis for disagreeing is sufficient to show that the sentence is not justified. Alternatively, one could also imagine Albertine responding, "But I know that Charles lies. He told Robert that he spent last year in Bangkok, when in fact he was in Manila." Perhaps for these two this fact outweighs what Charles says. But let's suppose that their conversation continues in the first way we imagined.

Given all we know about the situation so far, can we imagine that Albertine and Odette still might mean different things by the sentence "Charles lives on Broadway"? If so, we must think that the tale can be coherently continued as follows. Albertine agrees that her reasons for asserting the sentence are not enough to justify it. But, she says, "Nevertheless Charles does live on Broadway." Now, Odette, in astonishment, asks her, "Why? Didn't we just agree that what you're going on is not enough to show that Charles lives on Broadway?" Albertine says, in return, "Yes, we agreed on that. And I still do agree with it. I don't have sufficient reason to say that Charles lives on Broadway. But he does. You see, I don't mean the same thing as you do when I say, 'Charles lives on Broadway'."

Our (or my, at any rate) intuition is that if we left it at this point, there is something strange or unsatisfactory about the story. Don't we sense that some further plot development is called for if the story is to make sense? How would Odette respond to the last thing that Albertine said? Wouldn't she naturally ask, "What *do* you mean, then, when you say "Charles lives on Broadway"? Are we talking about the same Charles? Or maybe you mean something else from what I do by 'Broadway'? Or perhaps we don't mean the same thing by 'live'?" What would Albertine say? Suppose she said, "Oh, well, by 'live' I mean projecting oneself into one's existential horizon, and you, I think, mean merely physically occupying a dwelling." How would the conversation proceed from here? We know that Albertine and Odette agree completely in what they take to justify the things they say. So wouldn't Odette say something like, "I see. Well, if you mean *that* by 'live', then, of course I agree that Charles lives on Broadway. It is only on Broadway heading uptown, before it crosses Amsterdam, that he projects himself into his existential horizon. But look, we are not really disagreeing, then, are we? We're both right. In fact, don't you agree that if, by 'live', you mean what I do, then you would be wrong and I would be right?"

The story now makes sense, but it no longer presents the possibility of these two people assigning different meanings to a sentence while agreeing completely about what would justify it. I think that if we tried to imagine Albertine giving any other natural answer to Odette's request to specify her meaning, what we would end up imagining would also fail to be this putative possibility.

Why do we demand this further continuation of the story? Suppose Albertine hadn't said that she didn't mean the same thing as Odette does by the sentence "Charles lives on Broadway". That is, suppose Albertine simply agreed that she didn't have sufficient reason to say that Charles lives on Broadway, but also insisted that he does, without giving any grounds. Doesn't this amount to a concession that Odette's grounds for disagreement gives Albertine reasons for her to withdraw her assertion? So, unless she has some basis for doubting Odette's grounds, or can give a reason for her assertion that outweighs Odette's basis for doubting it, surely she *ought to* withdraw her assertion. And if she doesn't, wouldn't she just be unreasonable? That is, wouldn't we think that she's acting irrationally?

Now, in the situation we have been imagining, we had arrived at a point when Odette's reason for not taking back her assertion is that she means something different from Albertine. Our sense that this is not a satisfactory ending derives from the following fact. We don't think that, simply by *saying* that she has a different meaning in mind, Albertine succeeds in having a different meaning in mind. Maybe she says this just to avoid having to withdraw her assertion. What distinguishes the case in which Albertine does have a different meaning in mind from the one in which she doesn't? It is hard to see what the distinction could amount to except something like this: Albertine genuinely means something different only if she can make explicit how the difference in meaning that she cites counts as a reason for the assertion. So that's why our imaginings ran aground: we have been assuming that Albertine and Odette agree completely in what they count as reasons for and against their assertions.

The moral of this story is the following general claim:

> Whenever there is an apparent disagreement over an assertion made with a sentence, *S*, between two individuals, *A* and *B*, who would count exactly the same arguments as justifying such an assertion, they will be able, in the course of an analysis of their disagreement, to arrive at the recognition that not both of the assertions are correct.

From this claim it follows that *A* and *B* cannot attach different meanings to *S*. For, if they did, the apparent dispute between them would be merely verbal, and they would be able to come to recognize that both of them are justified in their attitudes toward the assertion made with *S*, contradicting the general claim.

All this shows something about the interrelations among our concepts of meaning, justification, and assertion. First, a difference in meaning could result in a difference in *being justified*. In other words, a difference in meaning provides a standard of being justified. But second, since assertion is a kind of rational action, being justified in an assertion is to be assessed in terms of the standards set by what we would count as justifications of it—otherwise we can make no sense of the idea that *we have reasons* for what we say. So, finally, we have difficulty forming a coherent conception of a difference in meaning in the absence of a difference in what is counted as justification, where the coherence in question is coherence with our practice of assertion.

As I pointed out at the outset, the problem that this argument raises has nothing to do with the question of how we can know what another speaker means by her words. The present problem is rather that, given our practice of assertion, we can't form a coherent conception of this meaning if we tried to divorce it from justification.

15.4 The Rejection of the Bivalence Argument

What does this conclusion about meaning have to do with classical logic? Well, let's first apply this account of manifestation to truth conditions. What we get is this. If two people attach distinct truth conditions to a sentence, then there must be a difference in what they count as justifications of assertions made with that sentence. One might think of it this way. A difference in truth conditions requires a difference in *how* the obtaining of those conditions can be recognized.

This already moves some distance from the underlying intuition of realism, which insists that there are no epistemic constraints on truth conditions. This distance doesn't amount to a rejection of realism. We need a further argument against a conception of truth conditions based on realism. The one I'll sketch is is one of those slippery burden of proof arguments. According to realism, our statements can express truth conditions whose obtaining or otherwise are not recognizable by us. How, then, are such conditions different from conditions whose obtaining *are* recognizable by us? Suppose a realist philosopher, by coincidence also called Odette, associates realist truth conditions with an undecidable statement, and an anti-realist philosopher, again coincidentally called Albertine, associates anti-realist truth condition with that very sentence. According to our second anti-realism, there must be some difference in what these two would count as justifications of this sentence. What could that difference be? It would have to amount to this. Odette must accept that she has some ways of recognizing the obtaining of truth conditions that precisely can obtain without being recognizable in any way by her or any of the rest of us. Is this idea coherent? Can there really be such ways? Questions like these wouldn't arise for Albertine, who believes that any condition she describes with her statements is a condition that she could recognize as justifying her asserting these statements.

Mathematical justification is proof. In the remainder of this story, we will focus on anti-realism about mathematics. Here the conclusion of the last argument implies that the truth conditions of mathematical statements are identical to their conditions of proof. The question we will now investigate is whether this claim has any consequences for deductive reasoning in mathematics.

Before addressing this question, it's worthwhile saying a bit more about what exactly the question is. Dummett explicitly models his general semantic anti-realism on (what he construes as) the intuitionists' rejection of classical logic in mathematical reasoning. So, there is an understandable tendency to equate semantic realism with classical logic, and semantic anti-realism with a rejection of classical logic. Someone in the grip of such a picture about the relation between anti-realism and classical logic is likely to suppose that an anti-realist and a classical logician must attach different meanings to the word 'true', the disjunction sign, and to the existential quantifier. To understand the question that we're about to tackle, one has first to reject all such presuppositions. One has to begin by leaving it open whether or not anti-realism entails a rejection of classical logic. The position from which we're now starting is anti-realism of the second variety described above. So, from now on, when I speak of 'the defender of classical logic,' I mean someone who accepts semantic anti-realism.

Let's start by reminding ourselves of the second episode in my story of the first anti-realism. There, classical logic fails from an anti-realist perspective because we don't know that undecidable statements are either true or false. The underlying presupposition of this bivalence argument is that the nature of deductive validity is fixed by semantic values. If there were no undecidable statements, then we would know that all statements are either true or false, and, therefore, that classical reasoning is logically valid. Another presupposition of the bivalence argument is the other component of the explanation of deductive validity: the meanings of the logical constants (partially) determine the truth conditions of statements in which they occur. In particular, the bivalence argument presupposes that from the meaning of "or" it follows that

> A statement of the form $\ulcorner p$ or $q \urcorner$ is true just in case at least one of p and q is true; and it is false just in case both p and q are false. (1)

But now let's come back to the story at hand. The second anti-realism that I have presented implies that the truth conditions of mathematical statements are not coherently distinguishable from their conditions of proof. So what, on this view, are the truth conditions determined by the meaning of "or"? Surely they are:

> We have (there is) a proof of a statement of the form $\ulcorner p$ or $q\urcorner$ just in case we have (there is) a proof of at least one of p and q; and we have (there is) a refutation of it just in case we have (there is) a refutation of both p and q. (2)

But, now, what does this claim say? We are, of course, focusing on mathematical statements, so the claim can be understood in two ways. Firstly, it can be understood as a claim about outright provability, as opposed to provability in a theory, i.e., about deductive justification from no premises or proof from logic alone. On this interpretation, the claim is: $\ulcorner p$ or $q\urcorner$ is provable from logic alone if and only if p is or q is. But what this amounts to is the claim that the system of purely logical inference applicable to mathematical statements satisfies the disjunction property. And that implies that this system of inferences cannot be classical logic. Secondly, we can interpret (2) to be about provability in a mathematical theory. In this case the claim it makes is: $\ulcorner p$ or $q\urcorner$ is provable from a theory if and only if p is or q is. This claim is satisfied by such theories as Heyting arithmetic, and by most intuitionistic mathematical theories. But it is not, in general, satisfied by classical mathematical theories.

Thus, if anti-realism is true of mathematics, then to accept that

> The meaning of 'or' in mathematical discourse determines that a disjunction is true just in case at least one disjunct is true (3)

is already to accept that no legitimate system of purely deductive inference applying to mathematics is classical. Put in this way, it would seem that a defender of classical logic should ask, why should I accept (3)? Note well that, as I stressed above, this classical logician is an anti-realist. So her question is not based on construing "true" in (3) as expressing realist truth rather than provability. The ground of her objection, rather, is this. The bivalence argument is supposed to be an argument against accepting classical reasoning by undermining the semantic justification of classical forms of reasoning such as the law of excluded middle. In order for it to be compelling for an anti-realist classical logician, the semantic principles assumed in the argument must not already rule out classical logic. Our anti-realist classical logician is committed to an identification of truth conditions with proof conditions. This identification has of course consequences for what counts as semantic facts or principles. But does it, by itself, entail that the proof conditions of disjunctions have to be given by (2)? It's hard to see that it does; why can't one hold that (2) states the proof conditions of all disjunctions except those in which q has the form $\ulcorner$not $p\urcorner$, in which case the disjunction is justified in all circumstances? Hence it would appear that the bivalence argument begs the question against the validity of classical logic.

Now, one might respond that (3) is surely no more than an expression of what we intuitively take to be the truth conditions of disjunctive statements. Hence, once we accept anti-realism, we are committed to the conclusion that the disjunction property holds of our system of deductive inference, on pain of giving up our intuitions about the meaning of disjunction. But then, on this line of thinking, although the bivalence argument is not question-begging, it is *unnecessary* for the anti-realist rejection of classical mathematical reasoning. Once anti-realism is accepted, we don't need the claim that undecidable mathematical statements are not guaranteed to be either true or false, in order to show that the validity of classical reasoning is unjustified. All we need is the intuitive claim that statements of the form $\ulcorner p$ or $q\urcorner$ are true just in case at least one of p and q is.

The bivalence argument is thus either question-begging or unnecessary. If the former, then, clearly, the rejection of bivalence fails to yield a cogent criticism of classical logic in mathematics. If the latter, then, again, the rejection of bivalence fails as the basis for the anti-realist criticism of classical logic. For, here the claim is that classical mathematics fails because of the truth conditions of disjunctions, and not because of facts about the semantic values of instances of the law of excluded middle.

A similar point can of course be made for the first-order existential quantifier. $\ulcorner(\exists x)Fx\urcorner$ is true if and only if $\ulcorner Fa\urcorner$ is true when a is assigned some member of the domain of quantification. Under the identification of truth conditions with justification conditions, this means that there is a proof of $\ulcorner(\exists x)Fx\urcorner$ under exactly the condition that there is a proof of $\ulcorner Fa\urcorner$, where a is singular term referring to some member of the domain of quantification. In the case of arithmetic, for instance, this amounts to the explicit definability property for natural numbers. This property, once again, rules out classical reasoning with existential quantification. And, again, the question is whether anti-realism's identification of truth with proof by itself is sufficient to entail this account of the proof conditions of existential quantification.

15.5 Proof-Theoretic Validity

But there is another way of considering the situation. If we accept the second anti-realism about mathematics, we are committed to identifying truth conditions with proof conditions. So far my story tells us that this identification implies that we have to think of our intuitive view of the truth conditions of disjunctions as a claim about the practice of mathematical justification. Now, it turns out that this claim conflicts with classical mathematical practice. That is, we have to change *something* if we are to make our intuitions about meaning cohere with our deductive practice. But what should we change? Our present anti-realism doesn't tell us. Prima facie, nothing in *anti-realism about mathematics* commits us to abandoning classical logic rather than the meaning that we think we attach to disjunction. Anti-realism implies that to understand a mathematical sentence, i.e., to grasp its truth condition, is to know what counts as a proof of it. But, at first sight, this by itself is consistent with the claim that, in grasping the sense of any sentence of the form, ⌜p or not p⌝, we count any condition whatsoever as one in which there is a proof of it, so that, in particular, we count it as proved even when neither p nor ⌜not p⌝ is.

Thus, we still have no criticism of classical logic on the basis of the second anti-realism. But this type of anti-realism, I will now argue, does imply a rejection of the interpretation of deductive validity in terms of semantic values that underwrites the bivalence argument. That interpretation is a way of spelling out what condition (V) requires, in terms of the semantic values of expressions. Can we consistently accept this account of (V), if we accept our second anti-realism? Let's ask how (V) must be understood from the vantage point of anti-realism about mathematics. That is, i.e., what *is* (V), if truth conditions are proof conditions? Surely, (V) becomes the following requirement on forms of inference:

If there are proofs of all statements that are instances of the premises, then there is a proof of the statement that is the corresponding instance of the conclusion. (V_{PC})

So, what follows from this? To begin with, it seems relatively uncontroversial that, in general, in giving a proof of a mathematical statement, we give a deductive argument for it. But surely it is no less controversial that it is not coherent to take a deductive argument to *justify* a sentence, unless one acknowledges that the forms of inference employed in the argument are valid. Hence we cannot in general coherently take a form of inference to satisfy (V_{PC}) without acknowledging the validity of at least *some* forms of inference.

But now we have a problem, if satisfying requirement (V_{PC}) is the analysis of the concept of deductive validity. If so, then whether any form of argument is deductively valid rests, ultimately, on whether it meets condition (V_{PC}). Now suppose we are trying to determine where a form of argument, A, is valid or not. In order to determine this, we have to be able to decide whether the rules of inference used in any set of arguments for the premises and the conclusion are valid. If A itself is one of these rules, then we have argued in a circle; otherwise we have to use (V_{PC}) to determine whether these rules are valid. Hence, the use of (V_{PC}) in determining whether we have reason to accept a form of inference will lead either to a circle or to an infinite regress. The conclusion is that our second anti-realism implies that deductive validity cannot be analyzed in terms of satisfying condition (V). Hence no criticism of classical logic starting from this anti-realism can be based on (V).

This conclusion doesn't rule out (V_{PC}) as *part* of the grounds for assessing the correctness of forms of inference. The reason is this. The argument we have just given is consistent with the supposition that there are forms or rules of inference that can be determined as valid on some basis *other than* their satisfying (V_{PC}). Let's call such forms of inference, if there are any, α-*rules*. Prawitz, following Gentzen, has taught us that, starting from general features of deductive reasoning, it is possible to define a notion of valid deductive argument on the basis of the assumption that α-rules are valid. A valid argument in this sense is valid in virtue of the α-rules already accepted as valid; call these *canonical arguments*. Now, restrict the justifications mentioned in (V_{PC}) to canonical arguments:

> Given canonical arguments for all statements that are an instance of the premise schemata, there exists a canonical argument for the statement that is the corresponding instance of the conclusion schema. (V_{CA})

(V_{CA}) can then be taken to be the condition for the validity of rules of inference other than the α-rules. Hence, if we can identify a set of α-rules, (V_{CA}) might provide an analysis of logical consequence consistent with the second anti-realism.

Of course this is all supposition. We have a lot of questions to answer. The main one, of course, is: what are these α-rules? What grounds, apart from satisfying (V_{CA}), could there be for determining whether a rule of inference is valid? What sorts of properties qualify a rule of inference as an α-rule? What rules of inference have these properties? We should also ask, is there only one type of rule of inference that can be an α-rule? But I suppose this latter question is not as pressing for our present topic.

An obvious suggestion is this: in arriving at the interpretation of validity in terms of semantic values, we used the meanings of the logical constants to give truth conditions of logically complex sentences in terms of the semantic values of their sub-sentences. Since the second anti-realism identifies truth with proof, this idea must be understood as the claim that the meanings of the logical constants give the proof conditions of logically complex sentences in terms of the proof conditions of their sub-sentences. But these proof conditions state when a logically complex sentence can be proved from its sub-sentences, and hence are claims about the validity of certain forms of argument. Hence, we can take these forms of inference to be α-rules. That is to say, a rule of inference is an α-rule if its acceptance as valid is required for someone to count as knowing the meaning of a logical constant.

An analysis of logical consequence based on such a set of α-rules and (V_{CA}) differs from the one based on the semantic value interpretation of validity, since it is not the case that the validity of *every* form of inference is determined by the satisfaction of condition (V). Let's call the standard for evaluating deductive inference based on this analysis the *anti-realist criterion of validity*.

What I have just described is of course just a type of proof-theoretic justification of logical laws. Our conclusion at this point is that if there is to be a critique of classical logic on the basis of the second anti-realism, then it must be based on proof-theoretic properties of classical reasoning, and not on the failure of bivalence.

The foregoing also shows what an anti-realist assessment of logical laws has to accomplish, in order that one might use it to provide an anti-realist critique of classical reasoning in mathematics.

The notion of an α-rule is so far explained only in terms of what we are committed to in virtue of knowing the meanings of the logical constants. But nothing in this explanation tells us whether any currently accepted rule of inference is *not* an α-rule. And we clearly have to show that not every currently accepted form of inference in mathematics is an α-rule; otherwise all classical forms of inference would be valid.

This can't be done on the basis of our second anti-realism by itself. The reason is this. On this anti-realism, to know the meaning of a mathematical statement is to know under what circumstances we have a proof of it. Surely this is possible only by knowing what forms of inference are valid. But nothing in the second anti-realism rules out that knowledge of the meaning of a mathematical sentence requires acceptance of the validity of all the rules of inference generally accepted in mathematical reasoning. So an anti-realist critique of classical logic requires a separate argument against this view.

But ruling out this view doesn't get us that far. If we succeed in ruling it out, we can conclude that acceptance of some forms of inference is more fundamental to our grasp of meaning than acceptance of others. But which rules are more fundamental? For all that has been shown, the acceptance of double negation elimination is required for someone to count as knowing the meaning of the negation constant. If so, then classical logic is justified because its basic forms of inference are α-rules, embodying the meanings of the logical constants. Hence, in order for an anti-realist critique of classical logic to succeed, the notion of an α-rule has to be further developed to disqualify basic classical forms of reasoning as α-rules.[3]

Finally, if these two problems can be resolved, it remains for the anti-realist to show that classical reasoning is not valid by the standards of the proof theoretic account of logical consequence.

Acknowledgements I would like to thank an anonymous referee, the editors, and Dr. Giuseppe Primero for their comments, advice, and help.

References

1. AA. VV. 2006. "Special Issue on Proof-Theoretic Semantics." *Synthese* 148.
2. Beaver, B. I. 1997. "Presupposition." In *The Handbook of Logic and Language*, edited by J. van Benthem and A. ter Meulen. Cambridge, MA: MIT Press.
3. Shieh, S. forthcoming. "Truth, Objectivity, and Realism." In *The Oxford Handbook of Truth*, edited by M. Glanzberg. Oxford: Oxford University Press.

[3] In particular, one needs an account of the identity conditions of formulations of α-rules. See, in particular the contributions by Prawitz, Contu, Tait, and Usberti in [1].

Chapter 16
A Double Diamond of Judgement

A Perspective on the Development of Logical Theory 1800–2000

Göran Sundholm

16.1 Introduction

In today's logic, *propositions* occupy the centre stage. They may be, and often are, viewed platonistically as independent entitles that are denizens of a 'Third Realm', next to those originating in either the material realm or the psychic one, perhaps construed as functions from 'possible worlds' to truth-values in a so called possible worlds semantics, but also linguistically, after the fashion of medieval logic, where a proposition is not a Platonist object in the third realm, but either mental or linguistic. Formalistically they are taken as well-formed formulae, 'WFF's', that is, as strings in a freely generated semi-group over a finite alphabet. Propositions are held to be the primary bearers of truth, both from the perspective of Platonist content, for instance, by Frege and Bolzano, as well as from the *sententialist* perspective of ontologically more parsimonious logico-philosophers such as Quine. Also under their formalistic construal as WFF's does it hold good that propositions serve as *truth-bearers*, as witnessed by Tarski's famous 'definition of truth' in *Der Wahrheitsbegriff*, where the recursive definition parallels the steps in the generation of the semi-group of strings.

Concomitant to their role as truth-bearers, propositions also serve as relata in the central notion of current logic, namely, that of (logical) *consequence*. Realists in the philosophy of logic customarily explain both the truth of propositions, as well as the consequence relation among them, in an ontological fashion, deploying (the obtaining of) states of affairs as a key-concept—be it further reduced model-theoretically in terms of satisfaction in set-theoretic structures or not—in terms of which other notions are then explained. This way of proceeding, which, under the influence of Tarskian model-theoretic semantics, has become all pervasive, has served to drive out virtually all epistemological concerns from the ambit of logic. Thus, for instance, it is common to deal with the epistemological concerns of logic only in terms of the *semantics* of WFF's such as

G. Sundholm (✉)
Institute for Philosophy, Leiden University, Leiden, The Netherlands
e-mail: goran.sundholm@gmail.com

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_16,

$K_a p$ – agent a knows proposition p,

where K is a suitable two-place propositional operator, taking an agent and a proposition, thereby yielding a novel proposition. This manner of proceeding has imposed too tight a straightjacket on the bond between logic and epistemology, a straightjacket, though, that was not present in the traditional logical picture that we meet around 1800, say, in Kant's *Jäsche Logic*.

16.2 Judgement and Inference: The Traditional Picture

Traditionally the crucial notion of logic was not that of a proposition (in the above, modern sense of an independent entity that serves as truth-bearer) but that of *judgement*. In the then prevalent psychologistic paradigm a judgement was primarily seen as an inner mental act that, nevertheless, could be exteriorised in an act of *assertion*, in which case the assertion made would be the judgement made, that is, the external correlate to the inner object, or product, of the act of judgement. A declarative sentence S can be recognised as such, it is well known from elementary textbooks in logic, by means of a criterion involving the appositeness of the question:

Is S true?

An equally simple—but surprisingly unfamiliar—consideration yields a criterion for identifying assertoric uses of declarative sentences. An utterance of a declarative is assertoric when the utterer is at fault unless he cannot provide an answer to the question:

How do you know that S is true?

Thus, as revealed by, for instance, Moore's paradoxical assertion of

It is raining, but I do not believe it,

assertions made contain implicit claims to knowledge of the truth of the proposition expressed.

These and other epistemic aspects are now largely lost to logic. In particular, their neglect has led to the expulsion of the epistemic notion of *inference*, that is, a passage from premiss-*judgement*(s) (made) to conclusion judgement, which has been replaced with that of (logical) *consequence*, that is, a relation between antecedent and consequent *propositions*. An (act of) inference is a *discursive* act of judgement that is effected on the basis of previous judgements made:

$$\frac{\begin{array}{cccc} | & | & & | \\ J_1 & J_1 & \ldots & J_k \end{array}}{J}$$

where the vertical bars above the judgements $J_1, J_2, \ldots, J_k$ represent acts that yield, respectively, the judgements made in question. The judgement made, that is, the object/product of the act of judgement, traditionally is of the subject/copula/predicate form [S is P], where S and P are terms, since an act of judgement is an act of composition (or division) with respect to two terms that have been previously obtained as results of two acts of simple apprehension. The traditional picture accordingly involves three interlinked notions, to wit those of *Term*, *Judgement*, and *Inference*:

THE TRADITIONAL STRUCTURE OF LOGIC[1]

OPERATION OF THE INTELLECT	MENTAL PRODUCT	EXTERNAL SIGN
Simple Apprehension	Concept, Idea, (Mental) Terms	(Written, Spoken) Term
Judging (Composition/Division of two terms)	Judgement (made), (Mental) Proposition: S is P	Assertion (made), (Written/Spoken) Proposition
Reasoning (Inferring)	(Mental) Inference	(Written/Spoken) Inference, Reasoning

The beauty and power of this traditional scheme ensured that it stood virtually unchallenged for (even more than) 2000 years:

> [S]ince Aristotle, [Logic] has not had to retreat a single step. Also remarkable is that it has not been able to take a single step forward, and thus to all appearance is closed and perfect.[2]

16.3 The Great Bohemian: Unary Judgement

One might take exception to this dictum of Kant's, but the mere fact that it could be, and also was, straightforwardly made shows that it contains a lot of truth. One immense step forward in the development of logical theory was to be taken half a century later by a Bohemian priest, who was also a theologian, philosopher of religion, mathematician, logician and ethicist, and arguably the greatest philosopher of the nineteenth century, to wit, Bernard Bolzano.[3] The aim of his endeavours was to provide an all-encompassing *objectivist* framework for philosophy and mathematics. From my present point of vantage his decisive contribution was to replace the traditional [S(ubject) is P(redicate)] bipartite form of judgement with the unary form

proposition(in-itself) A is true.

[1] This diagram, which I have often used in previous writings, is reasonably standard, and based on one in [3]. His source, however, like that of most Neo-Thomists, is the *Ars Logica* by John of St. Thomas.

[2] Kant, *Kritik der reinen Vernunft*, B VII, my translation.

[3] See his *Wissenschaftslehre*. Note the spelling; contrary to an almost universal misapprehension, to which many writers on Bolzano, including myself, have fallen prey, Bolzano's first name does not contain an 'h'.

Bolzano's term for his notion of proposition was *Satz an sich: proposition in itself*. A *Vorstellung an sich*, that is, an idea-in-itself—is a part of a proposition-in-itself that is not a proposition-in-itself. Any proposition-in-itself A has the form [a has $\boldsymbol{B}$], or is equal to a proposition-in-itself of this form.[4] For instance, the proposition [No man is mortal] is equal to the proposition [the Idea *mortal man* does not have *Gegenständlichkeit*].[5] The proposition [a **has** $\boldsymbol{B}$] is true (in-itself) precisely when *a really* does have B. These two *an sich* notions, that is, propositions-in-themselves and their truth-in-itself, constitute the pivot around which Bolzano's logical universe revolves. In particular, Bolzano treats of the validity of inference by means of a reduction to a notion of (logical) consequence explained in terms of these two notions. Owing to the change in the form of judgement the general inference figure I:

$$\frac{J_1 \;\; J_1 \;\; \ldots \;\; J_k}{J}$$

is transformed into a more definite figure I':

$$\frac{A_1 \text{ is true } \; A_2 \text{ is true } \; \ldots \;\; A_k \text{ is true}}{C \text{ is true}}$$

A judgement made of the form [A is true] is correct (*richtig*), according to Bolzano, if A *really* is true. Similarly, an inference carried out according to the schema I' is valid if the proposition

$$A_1 \& A_2 \& \ldots \& A_k \supset C$$

is *logically* true, that is, is true under all variations of non-logical parts.[6] Judgement is an epistemic notion and so is inference, being primarily an act of judgement made from, or on the basis of, other known judgements. An inference figure (or, perhaps better *inference candidate*) I' is valid when it preserves knowability from premises to conclusion. A consequence, on the other hand, be it logical or not, is an *alethic* notion since it is ultimately concerned with the preservation of truth from antecedent propositions to consequent proposition. Bolzano now has two reductions of epistemic notions to alethic ones; firstly, a judgement [A is true] is correct (*richtig*) if the proposition really is a truth-in-itself, and, secondly, the inference candidate I' is valid when the matching implication is a logical truth (or equivalently when the consequence $A_1, A_2, \ldots, A_k \Rightarrow C$ holds logically, that is under all variations of non-logical parts). The joint effect of these two reductions is to

[4] For the sake of brevity, in the sequel I take the In-itself qualification as understood.

[5] An Idea is *Gegenständlich* when an object falls under it. Accordingly another equivalent would be [the Idea *mortal man* is not objectual]. [9] contains a beautiful presentation of Bolzano's work using *Gegenständlichkeit* as a key-concept.

[6] This is not exactly true; Bolzano imposes some further conditions on his consequences, for instance, that the antecedent propositions be compatible. However, at the level of abstraction at which I want to move, it is more than true enough.

give a beautifully smooth realist epistemology.[7] However, in my opinion it is fatally flawed: a blind judgement is rendered correct, and, according to Bolzano, it is an *Erkenntnis* (a piece of knowledge) simply in virtue of pertaining to a proposition that happens to be true, and, similarly, blind inference, irrespective of whether it preserves knowability from premise(s) to conclusion is rendered valid simply in virtue of a (logical) consequence-relation obtaining between propositions. In sum then, Bolzano introduced a versatile unary form of judgement and the very important notion of consequence among propositions (be it logical or not).[8] The logical form of his propositions-in-themselves, on the other hand, was a relatively clumsy one, being in essence little but the Aristotelian $S/is/P$ formed thrust into the proposition instead.

16.4 Brentano and an Alternative Unary Approach

Franz Brentano, another priest-philosopher, was responsible for an alternative revision of the bipartite Aristotelian form of judgement to a unary form. Whereas Bolzano's unary form was propositional, Brentano's logic is a term logic, and his unary form of judgement accordingly involves a general concept α: the judgements

α exists (IS), in symbols $\alpha+$

and

α does not exist (IS NOT), in symbols $\alpha-$

are true when something falls under the concept α, respectively when nothing falls under α. The usual Aristotelian categoric judgements and their use in syllogistic inference are then readily treated of.[9]

---

7 Indeed, the version offered by Moore and Russell in their early apostasy from British Hegelianism during the first decade of the twentieth century is much inferior to Bolzano version. The quality of Bolzano's treatment is enhanced by his arguing primarily against Kant, that is, the foremost exponent of epistemological idealism, whereas Moore and Russell have Bradley and Bosanquet as their target.

8 Consequence is a *three*-place relation in Bolzano, pertaining to antecedent and consequent propositions, as well as to (a collection of) Ideas occurring in these propositions at which places where the variation takes place, under which truth has to be preserved from antecedent(s) to consequent(s). The consequence holds logically when truth is preserved under all variations with respect to all non-logical Ideas. When variation takes place at fewer place the consequence in question will not be logical. it is a moot point whether Bolzano allows for the natural terminus of preservation of truth from antecedent to consequent with respect to variation in no places. The corresponding consequence relation $A \Rightarrow B$ between the propositions A and B holds (but, in general, not logically) when the implicational proposition $A \supset B$ is true (rather than logically true. When this is the case, the consequence holds logically, of course.).

9 Brentano's use of his [$\alpha+$] form of judgement is completely parallel to Bolzano's use of 'the idea-in-itself V has *Gegenständlichkeit*'. To my mind, in view of Brentano's vehement and slightly undignified protestations to the contrary (that were printed by Bergmann [1, pp. 307–08]), a direct influence from Bolzano seems quite likely.

16.5 Frege's Judgement: Truth Applied to Function/Argument Structure

In spite of its title, Frege's *Begriffsschrift* refines the propositional rather than the *Begriffs* version of the unary theory of judgement. Frege also uses the unary form judgement [proposition A is true], but his form of proposition is superior to Bolzano's clumsy [V_1 has V_2]. The *mathematician* Frege, at this stage, apparently in complete independence from Bolzano, employs instead the mathematical function-argument form [$P(a)$]. From my present point of vantage Frege's advance in logical theory with respect to Bolzano is a minute one. Bolzano had transformed the (revised) Aristotelian form of judgement [V_1 has V_2] into a form of proposition. This propositional form Frege replaces with the mathematical function/argument structure [$P(a)$]. Furthermore, Frege, in contradistinction to Bolzano, did not consider the notion of (logical) consequence between propositions, but was solely interested in inference from judgement(s) known to judgement made. Throughout his career Frege also insisted that truth is *undefinable*.

16.6 Cambridge Truth-Making

A decade after Frege, the first examples of so-called *truth-maker* analyses emerge in the works of Moore and Russell, during their apostasy from British Hegelianism.[10] A truth-maker analysis does offer an elucidation of truth, after the fashion of the correspondence theory of truth:

Proposition A is true $=$ Truth-Maker(A) exists.

Thus a truth-maker analysis demands for each proposition

(i) a suitable notion of truth-maker for A, and
(ii) a concomitant notion of existence applicable to the truth-maker(s) in question.

Moore opted for *facts* as truth-maker, whereas Russell's choice was that of *complex*. The foremost example of a truth-maker analysis was provided by Wittgenstein in his *Tractatus*. Here, the truth-making applies only to atomic, or *elementary* propositions. The truth-functional mode of generation for complex propositions then serves to determine the relevant truth-conditions in terms of this truth-maker given notion of truth with respect to elementary sentences. It is important to stress that for Wittgenstein a proposition is a symbol, that is, (meaningful) sign in use; his propositions are not denizens of the Third Realm. For Frege the proposition, or Thought, was the *Sinn* of the meaningful sentence, whereas for Wittgenstein the sense of the meaningful sentence (proposition) is as situation (*Sachlage*) in 'logical

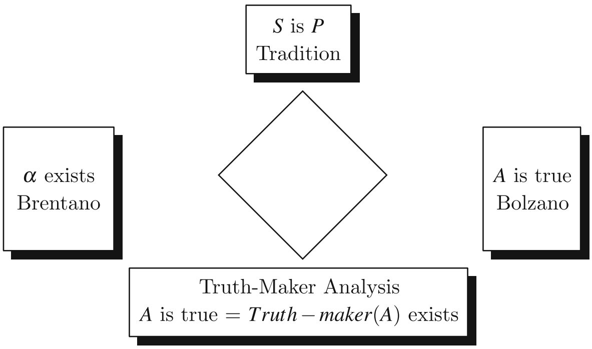
[10] The apt truth-maker terminology was introduced in [8].

space', that is, a directed partition of possible worlds into those compatible with and those incompatible with the situation in question.

An *elementary* proposition E presents (*stellt dar*) a state of affairs (*Sachverhalt*) S_E that may or may not obtain. When the presented state of affairs does obtain the proposition is true and otherwise false. Thus, in the Tractarian truth-maker analysis, the primary truth-bearer is the elementary proposition, and its candidate truth-maker is the presented state of affairs, whereas the relevant notion of existence is the ontological primitive of obtaining (*Bestehen*). The ensuing relation of truth-making, however, is then defined rather than primitive:

> The state of affairs S *makes true* the elementary proposition $E =_{def} E$ presents S and S obtains.[11]

The truth-maker analysis completes the '*first diamond of judgement*':



16.7 Constructivist Alternative: Proofs of Propositions

A decade after the *Tractatus*, an anti-realist version of truth-maker analysis was offered by Brouwer's pupil Arend Heyting in his proof-theoretical explanation of the intuitionistic, or *constructive*, notion of proposition. Here propositions are construed as sets of 'proofs'. Arguably, proof- or verification-object is a better terminology here. Previously, in the history of (epistemo)logic and mathematics, all proving ('*demonstration*') has taken place at the level of judgements (*theorems*). *Proof* (-object) *of a proposition* is a notion novel with and to intuitionism.[12] The traditional explanations run as follows:

[11] Connoisseurs of the *Tractatus* are requested to meditate on theses 4.022 and 5.542 at this point.

[12] Heyting's meaning explanations were meant to capture Brouwer's practice and were first offered in the early 1930s, e.g. [2]. The present streamlined version is based on the explanations offered in the constructive type theory of Per Martin-Löf [4–6].

STREAMLINED PROOF-EXPLANATIONS

$\bot$	There are no canonical proofs for $\bot$.[a]
&	When a is a proof for A and b is a proof for B, $< a, b >$ is a canonical proof for $A\&B$.
$\vee$	When a is a proof for A, $i(a)$ is a canonical proof for $A \vee B$. When b is a proof for B, $j(a)$ is a canonical proof for $A \vee B$.
$\supset$	When b is a proof for B, provided x is a proof for A, $\lambda x.b$ is a canonical proof for $A \supset B$.[b]
$\forall$	When D is a set, P is a proposition, when $x \in D$, and b is a proof for P, when $x \in D$, $\lambda x.b$ is a canonical proof for $(\forall x \in D)P$.
$\exists$	When D is a set, $a \in D$, P is a proposition, when $x \in D$, and b is a proof for $P[a/x]$, $< a, b >$ is a canonical proof for $(\exists x \in D)P$.

[a] The constructive general truth-maker analysis for atomic sentences remains to be given; presumably it would require a constructive theory of moments.
[b] The fully explicit form of this $(\supset I)$ rule is: $\supset I(A,B,(x)b) : A \supset B$ provided that $b{:}B(x{:}A)$, where $(x)b$ is a function obtained by ("λ"-) abstraction from the (dependent) proof-object b.

One must here note that the constructivist account makes use of an intentionalized notion of object; in particular, a proof-set Proof(A) is explained in terms of how canonical proofs of, or for, the proposition A may be formed and a proof is any object that is equal to—*evaluates to*—a canonical proof.

In order to complete the constructivist truth-maker analysis we also need a suitable notion of existence:

Proposition A is true $=$ $Proof(A)$ exists.

Clearly, on pain of a vicious regress, the appropriate notion of existence cannot be that of the constructive existential quantifier given by the final clause in the table above. Instead, what is involved here is the constructive notion of existence found in the mathematical practice of Brouwer and Kronecker, but to the best of my knowledge first formulated explicitly by Hermann Weyl.[13] Thus, when α is a type ('general concept')

α exists

is a *judgement*. Its assertion-conditions are given by the rule:

$$\frac{a \text{ is an } \alpha}{\alpha \text{ exists}}$$

Accordingly one is entitled to assert that there is an α only after having exhibited one. Here, at a surprisingly late stage, the constructive notion of existence makes

[13] Cf. Weyl's [12] notion of *Urteilsabstrakt*. In [10], I speculate upon Pfänder and Schlick as possible sources for Weyl, whereas the smooth version offered here goes back to Martin-Löf [7].

the choice of logic definite: in the absence of a proof-object for $A \vee \neg A$, we are not entitled to draw upon the (instance in question of the) law of excluded middle. If, however, one is prepared to employ a liberal notion of existence that admits also of non-constructive modes of reasoning it is very easy to give a proof-object for $A \vee \neg A$. The proof-explanation of the logical constants is logically neutral in much the same way that Tarski's truth definition is; the relevant features of truth are inherited from the logical means available in the metatheory.[14]

It is a remarkable feature of the intuitionistic theory that one returns to the original, traditional [S is P] form of judgement, while incorporating both the Bolzano- and Brentao-forms [proposition A is true] respectively [concept α exists]. The Bolzano- and Brentano-forms are joined in the truth maker-analysis

Proposition A is true = $\mathit{Truth} - \mathit{maker}(A)$ exists.

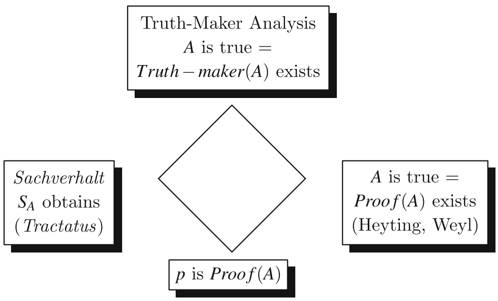
On the constructive analysis, where

$\mathit{Truth} - \mathit{maker}(A) = \mathit{Proof}(A)$

and existence is *constructive* existence, the traditional [S is P] form returns, namely

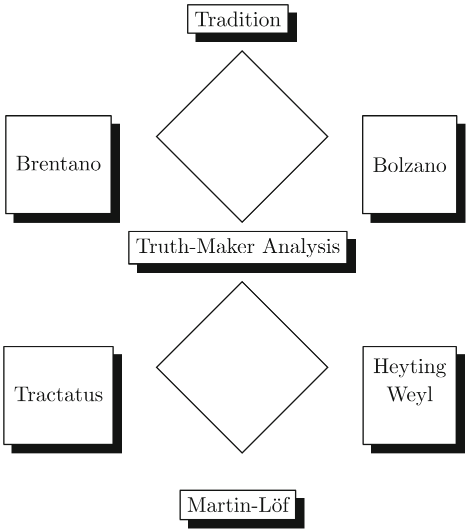
p is $\mathit{Proof}(A)$.

With this the '*second diamond of judgement*' has been completed:



The development of judgemental theory over the past two centuries is accordingly cast in the mould of a '*Double Diamond of Judgement*' with the Truth-Maker analysis in the pivotal role:

[14] See [11], i.e. the other half of my Geneva 2002 lecture, deals with this matter in considerable detail.



Acknowledgements This text is (a part of) an invited lecture at the international workshop 'Formal Ontology', Geneva, November 2002. The material was also presented as an 'Alumni Lecture' at the Dutch Graduate School in Logic, Amsterdam, March 6, 2003. I am indebted to organisers and participants alike, and to Kevin Mulligan in particular.

References

1. Bergmann, S. H. 1968. "Bolzano and Brentano." *Archief für Geschichte der Philosophie* 48:306–11.
2. Heyting, A. 1931. "Die intuitionistische Grundlegung der Mathematik." *Erkenntnis* 2:106–15. English translation in *Philosophy of Mathematics*, edited by P. Benacerraf and H. Putnam (2nd Edition).
3. Maritain, J. 1946. *An Introduction to Logic*. London: Sheed and Ward.
4. Martin-Löf, P. 1982. "Constructive Mathematics and Computer Programming." In *Logic, Methodology and Philosophy of Science VI*, Hannover 1979, edited by L. J. Cohen et al., 153–75. Amsterdam: North-Holland.
5. Martin-Löf, P. 1983. "On the Meanings of the Logical Constants and the Justifications of the Logical Laws." Lectures delivered in Sienna, first distributed in 1985, and printed in *Nordic Journal of Philosophical Logic* I:1 (1996): 11–60. Electronically available at http://www.hf.uio.no/filosofi/njpl/.
6. Martin-Löf, P. 1984. *Intuitionistic Type Theory*. (Notes by Giovanni Sambin on Lectures Given at Padua, June 1980.) Naples: Bibliopolis.
7. Martin-Löf, P. 1994. "Analytic and Synthetic Judgements in Type Theory." In *Kant and Contemporary Epistemology*, edited by P. Parrini, 87–99. Dordrecht: Kluwer.
8. Mulligan, K., P. Simons, and B. Smith. 1984. "Truth-Makers." *Philosophy and Phenomenological Research* 44:287–321.
9. Sebestik, J. 1992. "The Construction of Bolzano's Logical System." In *Bolzano's Wissenschaftslehre 1837–1987*, edited by Leo S. Olschki, 163–77. Firenze: International Workshop Firenze, seetembre 16–19, 1987.

10. Sundholm, B. G. 1994. “Existence, Proof and Truth-Making: A Perspective on the Intuitionistic Conception of Truth.” *TOPOI* 13:117–26.
11. Sundholm, B. G. 2004. “The Proof-Explanation of Logical Constants Is Logically Neutral.” *Revue Internationale de Philosophie* 58(4 [special issue on intuitionism edited by Michel Bourdeau]:401–10.
12. Weyl, H. 1921, “Über die neue Grundlagenkrise der Mathematik.” *Mathematische Zeitschrift* 10:39–79. English translation in *From Brouwer to Hilbert*, edited by Paolo Mancosu, 86–118. Oxford: Oxford University Press, 1998.

Chapter 17
Stable Philosophical Systems and Radical Anti-realism

Joseph Vidal-Rosset

17.1 Philosophical Systems and Philosophy of Logic

17.1.1 Vuillemin's Classification

Roughly, there are two main ways of developing what one calls 'a philosophy of logic'. The first way consists in starting from the actual scientific knowledge of logic, with the goal of presenting and, if possible, of solving problems linked with language and knowledge in general. Such questions could be for example, 'how to define logical consequence?', 'how to define negation?', 'how can we make use of the deduction of an absurdity?', 'how can we compare a logical system to another one?', and so on. It is obvious that this first way has at least this merit: making efforts to stick to the scientific reality of formal logic; but it has also the principal disadvantage of lack of unity and lack of structure in the philosophical inquiry. Imagine a mathematician with a weak philosophical culture, who is wondering about the logical points that a book of philosophy of logic must deal with, and you guess that this first road for the philosophy of logic leads to eclectism.

The second way to develop a philosophy of logic is based on a philosophy of knowledge in which philosophy of logic is embodied. Of course, two questions arise naturally when one sees that possibility. The first question is how to define the philosophy of knowledge, the second one is how to analyse, from a philosophical point of view, the role of logic within knowledge. Before answering this couple of questions, I am going to explain why I think that this second way is the right one for philosophy of logic. It is my contention that there is only one logic in the broad meaning of this word, I mean mathematical logic. (I claim the right to dismiss the existence of another logic like, for example, the so-called 'dialectical logic' expressed by the very obscure Hegelian jargon.) But it is also an unquestionable fact that, since the early twentieth century, formal logic has been developped and it has reached by now a very high level of specialization. Consequently, must

J. Vidal-Rosset (✉)
Archives Poincaré, UMR 7117, Nancy, France
e-mail: joseph.vidal-rosset@univ-nancy2.fr

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_17,

we choose between the risk of doing philosophy of logic without enough skills in understanding philosophical problems, or doing it without being, competent enough on the technical features of mathematical logic? Ideally, I believe that we must start from a clear understanding of the key philosophical questions of philosophy of knowledge in order to try and understand how mathematical logic is a tool for a potential resolution of these issues, or at least that can perhaps shed new light on these questions.

The main reason can be understood by an analogy. Every one can admit that logic is just a part of knowledge. So if philosophy of knowledge is a rational task, it seems rational also to conceive philosophy of logic as a part of philosophy of knowledge. If we come back to our mathematician knowing little philosophy, she is surely capable to see some philosophical points in her knowledge of mathematical logic. But these philosophical elements will need in turn to be related to the more basic and more simple issues of the philosophy of knowledge, and the elements that brings philosophy of logic will need also to be related with the other fields of philosophy of knowledge. Without the philosophical concepts, the mathematician will be lost in philosophy, like a mason without architect. But without the technical skills of a good mason, no architect can hope to build solidly. My analogy of course does not implies in the least contempt for mathematical knowledge, on the contrary, it just establishes the place taken by knowledge in the academic discipline which is always conceived as a meta-knowledge, i.e. a general knowledge about knowledges. The analogy would work in the same way, with philosophy of law as a part of philosophy of politics. *Mutatis mutandis*, the lawyer would take the place of the mathematician in such a case.

Our first job is then to define philosophy of knowledge before wondering about issues of philosophy of logic. In [18] and mainly in [17], books unfortunately not very well known, Vuillemin has shed a new light on the nature of the main classes of philosophical systems. It is not the place here to give a detailed presentation of Vuillemin's classification, but only to use it in order to clarify both the relation between philosophy of knowledge and philosophy of logic and, more specifically, to understand precisely the philosophical position called 'strict finitism'.

A philosophical system is defined in [16] as a 'system of integral reality',[1] i.e. a 'systematic ontology'.[2] Classes of philosophical systems are defined by different privileges given to different types of elementary sentences which are used to communicate perceptions. These differences in choices involve disagreements about what is reality and on what is only appearance, which explains why philosophy is essentially polemical.[3] The sharp division existing between the group of sentences that Vuillemin calls 'the dogmatic series' (i.e. denoting sentences which pay no

[1] See [16, p. 286].

[2] [17, p. 166].

[3] See [17, p. 113]:

> Consequently, as applied to ontology, axiomatics inevitably produces pluralism and disagreement. Indeed, philosophical reason is born and lives in contest.

regard to the speaker's subjectivity) and the group of 'subjective series' to which belong the judgments of method and the judgments of appearance, defines a division between two group of classes of philosophical systems: the dogmatic systems and the systems of examination. The adjective 'dogmatic' has no pejorative accent here. A 'dogmatic' philosophy, in Vuillemin's meaning of this word, denotes only a philosophy accepting as true the proposition 2.161 of [19]:

> In the picture and the pictured there must be something identical in order that the one can be a picture of the other at all.

Of course, dogmatic systems disagree about the linguistic forms that can be accepted as real, i.e., as belonging both to the picture (our language) and to the world. From the point of view of Platonism (i.e. 'realism' in Vuillemin's terminology), the denotation of universals is Ideas and Ideas are real. Representing the class of conceptualism, Aristotle rejects the platonic hypostasis: the universals are, from his conceptualist point of view, the result of abstractions (made by the mind) of qualities that are in things. The nominalist goes further. The nominalist critics is based on the following key thesis: universals are not even concepts, but only words, and Platonist school like Aristotelian philosophers are making the same error of confusing things (which are atoms or events) with words. From Platonism to nominalism, via conceptualism, the ontological landscape is more and more desert.

But independently of their ontological disagreements, these three classes of philosophical systems accept implicitly what Vuillemin calls 'dogmatism' and what Tennant describes as the realist conception of truth, illustrated by the following quotation from Russell[4]:

> On what may be called the realist view of truth, there are 'facts', and there are sentences related to these facts in ways which make the sentence true or false, *quite independently of any way of deciding the alternative*.

It is precisely that point, shared by the three main classes of dogmatic philosophical systems, namely the semantic realism, which is both denied by intuitionism and scepticism. For the former the true *never* transcends justification; for the latter, the true is just an illusion, and the philosophical work of the sceptic is to determine the degree of that illusion. I will not develop the case of scepticism in this paper. I aim only at showing why Vuillemin's understanding of what can be called 'philosophical intuitionism' explains also what animates the changes that intuitionist logician have made in logic, and how stable such a philosophical system is. The following quotation sums up the intuitionist argument[5]:

> From the intuitionist point of view, the whole dogmatic series of classes of systems that assume objective existence independently of any activity of the knowers offers just so many illusions: a philosophical illusion being raised when the role of the subject is juggled away behind the veil of an object provided with an alleged autonomy. As soon as a philosopher abandons reflection and the possibility of experience, he gives in to illusions, which each intuitionist system describes in reference to the method it acknowledges as its standard. [...] In all cases, however, illusions depend on the same excess of our infinite aspirations

[4] [10, p. 245], quoted by Tennant [13, p. 1]
[5] [17, p. 125].

> over our bounded power and on the same blind trust in the operations of logic conceived as an organon of philosophy instead of being confined to the role of is canon. Doubt is to be recommended every time the object of knowledge does not meet adequate procedures of construction.

The origin of intuitionism, from Vuillemin's point of view, lies already in philosophical systems like Epicurus', Descartes' and Kant's, where the determination of truth is subordinated to the method of decision (which can be defined in various ways): in many ways these three important philosophical systems lead to the general key thesis according to which truth is provability. But it is not my goal to develop here this historical-philosophical interpretation, masterly explained by Vuillemin in his book. It is time now to see how a philosophy of logic develops itself inside a more general philosophy of knowledge.

17.1.2 What Is a Stable Philosophical System?

There is a deep reason why every rational philosophy of knowledge is forced to develop a philosophy of logic: logic being the kernel of every reasoning and every theory, it offers to philosophy a privileged field of inquiry, because, by nature, philosophy seeks universality. But one must note that there is no disagreement between dogmatic systems neither on the use of the Principle of Bivalence in Mathematics nor on the universal validity of the Law of Excluded Middle (LEM). That probably explains correctly why the Tarskian interpretation for which a mathematical formula is either *true or false* can often be understood as *philosophically neutral*. But as soon as the Tarskian interpretation is adopted also with the Principle of Bivalence, it is not fair to interpret it as purely neutral from a philosophical point of view, because it seems hardly disputable that it recalls the semantic realist thesis expressed in Russell's quotation above.

The change vis-à-vis the Tarskian interpretation of formulas has been made by the Brouwerian-Heyting-Kolmogorov (BHK) interpretation of logical connectives which is a well-known expression of this intutionistic claim according to which mathematical truth is provability. Because in the BHK interpretation, a formula is true if and only if it has a proof, that involves of course a modification of the truth-conditions of formulas such $(p \vee q)$ or $(p \to q)$. While the classical logician concludes that $(p \vee \neg p)$ is a universal truth because of Bivalence and because of the truth-table of the disjunction, the intuitionist refuse to assert Bivalence and the rejects the universal validity of $(p \vee \neg p)$. The famous intuitionist refusal to assert LEM finds a clear explanation thanks to the BHK interpretation: the intuitionist will assert a disjunction *only if he is in a position to assert one of the disjuncts*; consequently it is easily understandable that there are many assertions that we are neither in position to assert, nor to assert their negation. These points are well known and have been clearly exposed in many textbooks.[6] The interesting point is

[6] See for example [1].

that intuitionistic logic provides of course a proof system to decide which formula or which deduction is intuitionistically valid. Because every intuitionistically valid formula (or every intuitionistically valid deduction) is also classically valid but not the converse, Intuitionistic logic is a part of Classical logic.

So, intuitionistic logic is born from a philosophical viewpoint about truth and knowledge (i.e., the semantic anti-realism) and has find a happy development inside the intuitionistic logical system, whose deductibility relation is a sub-relation of the classical one. That explains both why intuitionism is both foundation and reform of logic, when it is assumed by some philosophers as 'the right logic', i.e., the most basic and simplest logic at work in hypothetico-deductive sciences.[7] One of the most interesting feature of intuitionism is its stability, and one might even say its 'logical stability': because constructive mathematics is classical mathematics carried out with intuitionistic logic,[8] the intuitionist theory of the world can be conceived as entirely based on the intuitionistic logic which is in turn based on the semantic anti-realism requirement. Noticing that Veldman and de Swart have discovered intuitionistic completeness proofs for intuitionistic first-order logic that deletes the embarrassing fact that intuitionistic logic seemed to have a completeness proof only in classical logic, Tennant writes about the intuitionist[9]:

> He had attained what, it was complained, he needed: an account that was philosophically stable, in the sense that the justification given for one's choice of logic did not have to exploit logical resources (in the metalanguage) that lay outside the chosen system.

Stability has seldom been pointed out as a crucial quality for a philosophical system. Vuillemin in [16, 17] has also used this concept in a broader meaning than Tennant's. In Vuillemin's theory, a philosophical system is stable if and only if it uses the same basic assertion as principle of explanation, in all areas of its systematic picture of the reality. The following quotation of Vuillemin refers to Quine's philosophy, and, conversely, outlines its unstability[10]:

> A celebrated philosopher, for example, vindicate sets, things and perhaps events as the furniture of his world, regretting in other respects that he is not able to eliminate from the universe the extensional tracks of the ideas. The same philosopher, in morals, grants himself leave to build a more parsimonious construction since the ideas, reduced to sets, work in the theory of science but are idle in the theory of action. Such a system borrows one of its components from realism [i.e. Platonism]—but from an impoverished realism—, the other or the others from nominalism. Now, even were its highest principle, 'Being is being as set or thing', recognized as necessary and sufficient condition to account the whole of experience, it would, owing to its eclectic basis, reveal a kind of internal instability.

It is not my task here to wonder if the philosophical project of seeking to build, on few principles, a global vision of reality able to include all aspects of what

---

[7] See for example the work of Tennant [13]. I put aside here technical points around his Intuitionistic *Relevant* Logic, which is *the* right logic, according to him.

[8] See [2], quoted by Tennant in [14].

[9] [13, p. 306].

[10] [17, p. 133].

is in relationship with humanity, is still a reasonable task. My point is simply to notice that the concept of stability for a philosophical system is variable. But such a variability does not undermine the importance of stability, on the contrary: it puts the stress on the stability of intuitionism, which can be conceived as entirely based on intuitionistic logic. Clearly, Platonism is also stable, but its stability requires more than classical first order logic. If one accepts the use of the Quinean norm of ontological commitment, one must conclude that Platonism requires, in addition of the classical first order logic, the classical impredicative theory of sets, like the Zermelo-Fraenkel set theory, allowing quantification on abstract objects and on transcendental truths (i.e. truths that are beyond every calculus). Of course a parallel could be made by pointing out that the Platonist develops with the help of classical logic all the theory of classical mathematics, as the intuitionist develops the constructive mathematics with help of the intutionistic logic. But if the battle is confined on the ground of logic alone, in the restricted meaning given to the term of ‘logic’ by Quine, the Platonist who whishes to defend the specificity of his position will be in a tight spot, because classical logic and semantic realism are also accepted by logicians and philosophers, who are ‘dogmatic’ in Vuillemin’s meaning of the word, but anti-Platonist. These also accept the classical first order logic because of its ontological neutrality. This issue can be sketched by the following situation, described by Tennant[11]:

> The source of the dialectical problem here is the asymmetry involved: the reformist’s logic is a proper sub-logic of classical logic, not the other way round. If the reformist takes up the gauntlet to provide a philosophically stable account of his logic, then by the same token the classicist may only seek to persuade him of the soundness of the disputed classical principles by providing an argument based somehow only on principles which the reformist can accept. If the latter turns out to be impossible (as it would be if the reformist were right!), then so much the worse for the polite attempts to accommodate the classical disputant. There will not be any neutral point from which to adjudicate the dispute.

The argument suggested here by Tennant, can be generalized as a fair rule in philosophical disputes as follows. Provided that the principle (or the set of principles) adopted by the opponent do not lead directly to contradiction, the only way to contest his position is to find an argument which starts from the opponent’s principle(s) and lead logically to unacceptable conclusions for the opponent. We are going to see, in the second part of this paper, how it is important in a philosophical inquiry to determine to which class of philosophical system a philosophical argument belongs, in order to think about the possible strategies that could be at hand for philosophical disputes.

The next section deals with a contemporary philosophical position called ‘radical anti-realism’, which is based on strict finitism in the foundations of mathematics. The reader will judge if the attempt at determining the class of philosophical system to which radical anti-realism belongs, and at wondering if it is a stable philosophical position, is illuminating or not.

---

[11] [13, p. 307].

17.2 A Case of Philosophical Dispute: Strict Finitism vs. Intuitionism

17.2.1 The Contemporary Strict Finitist Argument

The following definitions are commonly accepted (I quote here [15]):

(A) Finitism is the thesis according to which mathematical objects and concepts have to be accessible to the mathematician in terms of constructions that can be performed.
(B) The strict finitist position rejects the potential infinite in the conception of construction of mathematical objects: if a procedure or algorithm will (provably) terminate at some moment in the future, then the outcome is accepted as constructable *if and only if* the length of computation is known as reachable. From a computational point of view, it means that an indefinite outcome is not acceptable, since all computational resources could have been used up before the outcome has been reached.

One must take care to distinguish precisely between strict finitism as previously defined, and Hilbert's finitism. The former pays attention to the feasibility of proofs when the latter 'roughly speaking, can be seen as a form of finitism on the meta-level (e.g., although mathematical theories can talk about infinite structures, still the proofs in such theories must have a finite length).' It must be noticed that the *non strict* finitist requires only finite length of proofs, which is simply a property of every effective proof. *Strict finitism demands more, requesting that every proof must be feasible in order to be considered as a proof.*

Recently, several arguments and directions of inquiry have been advanced by Dubucs and Marion in [4, 5], and in [8], arguing that *strict finitism* is a valuable position in philosophy of logic. The logical signification of the radical anti-realism has been also developed and analysed in details chapter 3 by D. Bonnay and M. Cozic, this volume.

My target consists now in determining to which class of philosophical systems (in Vuillemin's classification) the thesis of radical anti-realism belongs. Strict-finitism, which insists on the fact that every human is finite and on the concomitant rejection of the potential infinite, *cannot be* a genuine intutionistic position. Indeed, it is in fact, in history of philosophy, an old nominalistic refrain: one can believe in actual infinite or in potential infinite only if one forgets that human understanding is finite and that words and signs are a help when men are unable to manage too numerous informations.[12] The Radical Anti-realist would be in a stable philosophical position if and only if she could provide a foundation to her philosophy on a logic corresponding to her requirements of feasibility and strict finitism, and, on the contrary, she would be in unstable position if, in order to base his philosophical argument,

---

[12] See, for example, Spinoza, *Ethics*, II,prop. XVIII, scol.; prop. XL, scol. I.

she needed to make use of concepts or theories that do not fit with her point of view, leaving for example the door open to the potential infinite.

To judge about the stability or the unstability of radical anti-realism, let me sketch what is the radical anti-realist philosophical argument based on a so called 'strict finitist thesis' related to complexity theory.

(1) **The Radical Argument itself**. I quote Dubucs [4, p. 214]:

> The thesis I am willing to defend (following, among others, Wright (1987)) is that the way traditional anti-realism proceeds to characterize these assertability-conditions is in any case inadequate, and that it would be better to replace the Dummettian perspective of effectiveness *in principle* by that of *practical feasibility*.

The argument could be also labelled the 'Revolutionary Argument'[13]: the crux of the matter is that, from the point of view of the Radical Anti-realist, intuitionism *à la Dummett* is still too much dependant of idealizations that are not enough frankly different from Platonist theses: between ϕ as effectively provable but not feasibly provable, and ϕ as transcendent truth, the radical realist does not see much difference.

(2) **The search of a logic for the Radical Anti-realism**. Perhaps such a logic *could be* the Linear Logic. (See chapter 3 by D. Bonnay and M. Cozic, this volume.)

(3) **The polynomial time criteria** would be the finally the real logical base of the Radical Anti-realist thesis: its virtue lies in not appealing to human limitations but to the polynomial-time computability in its definition of feasibility. (See Marion [8]).

The following sub-sections are going to show what one can think of the hypothesis of finding the logic of the Radical Anti-realism, if this logic is the Linear Logic, or if it is the logic of P, i.e. the logic of problems decidable in polynomial time. Of course the developments about (2) and (3) must help to think about (1).

17.2.2 Linear Logic and Radical Anti-realism

Because Linear Logic is a complicated logical matter from a technical point of view, and because it does not seem to be considered, in the end, as being *the* logic of the Radical Anti-realist, I am going only to resume why it *could* appear as the correct logic from the Radical Anti-realist point of view, and I suggest readers wishing to know more that they read chapter 3 by D. Bonnay and M. Cozic, this volume. Proof-theoretically, Linear Logic derives from an analysis of classical sequent calculus in the absence of the structural rules of weakening and contraction. The originality of Linear Logic is to establish an interesting distinction between *persistent truths* and *ephemeral resources* and this point is one of explanations of its success in computer

---

[13] *Mutatis mutandis*, one finds an analogous argument in history of politics, when communists argued that socialist reformists did not go further enough.

science. This quote of D'Agostino et alii. [3] (in [9]) provides a clear explanation on that topic:

> Linear Logic completely rejects the 'vagueness' of traditional proof theory concerning the use and manipulation of assumptions in a deductive process. A 'proper' proof is one in which every assumption is used *exactly once*. If a particular assumption A can be used ad libitum, this had to be made explicit by prefixing it with the storage operator !. This means that the Contraction rule is not sound in Linear Logic, since it says, informally, that a proof of B from two occurrences of A can be turned into a proof of B from only one occurrence only of A. But this is impossible, unless A is one of those assumptions which can be used ad libitum, in which case we should prefix it with the storage operator.

The way in which Girard, who invented Linear Logic, has defined the enterprise of his logical system can lead one to believe that Linear Logic makes also Anti-realism more precise from a logical point of view[14]:

> Linear logic is a logic behind logic. More precisely, it provides a continuation of the constructivization that began with intuitionistic logic. The logic is a strong as the usual ones, i.e. intuitionistic logic can be translated into linear logic in a faithful way.

But because it is possible to translate into the language of Linear Logic both classical logic and intuitionistic logic, it appears that it is just *a fragment* of Linear Logic that could be *the logic* of the Radical Anti-realism. This fragment, as point out d'Agostino et alii. [3, p. 418] can be interpreted as an expression of Relevant Logic:

> The Contraction rule embodies the traditional (classical, intuitionistic and relevantist) careless approach which does not distinguish between using a formula once and using it any number of times. By disallowing Contraction, Linear Logic eliminates this residual degree of vagueness from traditional proof-theory: if a formula is to be used n times it must be assumed n times. In term of left-handed Gentzen systems, and according to this 'reductionist' point of view, the exponential-free fragment of Linear Logic can be seen as arising from the sequent system for **LR**, the distribution-free fragment of **R**, by removing contraction, all the definitions of the logical operators—as well as the distinction between multiplicative and additives—remaining the same.
>
> It is straightforward to translate the Linear Logic deductive policy into a stricter *criterion of use*: a Linear proof is a Relevant proof in which each formula is used *exactly once*.

But, unfortunately, the fact that 'the exponential-free fragment of Linear Logic can be seen as arising from the sequent system for **LR**' is not a good reason for choosing Linear Logic as *the* logic of the Radical Anti-realism. Tennant in [12, p. 7], points out that the propositional relevance logic **R** of Anderson and Belnap is undecidable, and that its well-known decidable fragment **LR** 'has an awesome complex decision problem: at best ESPACE hard, at worst space-hard in a function that is primitive recursive in the generalised Ackerman exponential.' Because it seems inappropriate to take a system whose complexity belongs to a class of complexity considered as *not feasible*, as the logic of feasibility, it seems more reasonable to throw Linear Logic overboard and to take another direction.

[14] [6].

17.2.3 The Feasibility Criteria: Polynomial Time Computabilty

In [5] Dubucs and Marion have assumed a criteria of feasibility whose merit is to be clearly based on the robust computer science. I cannot give a better account of the significance of this criteria than Marion himself proposes, thus I quote him[15]:

> In a joint paper with Dubucs, we suggested that one looks at the 'Karp–Cook Thesis' as a suitable alternative. According to it, a set of strings is feasibly computable if and only if it is 'polynomial-time computable'. Recall that a function $F(n)$ is said to be 'bounded from above' by another function $G(n)$ if for all n from a certain point on, $F(n)$ is no greater than $G(n)$, so that a function is 'polynomial-time computable' if it is bounded from above by a polynomial n^k, where n is function of the size of the input and k a constant. [...] There is no explicit appeal to human limitations here, so no clash with the above requisite, and no threat of semantic incoherence—as a matter of fact all problems linked with vagueness are bypassed. Furthermore, this approach to polynomial-time computability is linked rather with a finitist or Aristotelian ontology of infinities that are conceived of as unbounded processes as opposed to totalities; it clashes with a rather radical, and hard to swallow, rejection of large finite numbers, for which room was made in Gandy's Thesis.

There is a general agreement in computer science for considering the complexity class P, of problems recognizable in some polynomial amount of time, as being an *excellent mathematical wrapper of the class of feasible problems*. Any problem not in P is certainly not feasible. On the other hand, natural problems that have algorithms in P, tend to eventually have algorithms discovered for them that are actually feasible.[16]

Such a scientific general agreement seems to show that the Radical Anti-realist is in the right direction. But, it is impossible not to notice that she has changed her shoulder: the insistence on the complexity class P is certainly justified from the feasibility point of view, but the question of knowing if we are able to construct the logic of P *remains the main open problem* of the Descriptive Complexity Theory, i.e. to find a logic L, such as for every class of finite structures C, C is definable in L, if and only if C is decidable in deterministic polynomial time. Such a problem is another expression of the famous big problem of knowing if P $=$ NP, if one can prove that it is impossible to get the logic of P, then one proves also that P $\neq$ NP. Last, because one usually suspects that P $\neq$ NP, one could also suspects that it is not possible to construct the logic of P. We are going to see in conclusion that there are two possible attitudes for the Radical Anti-realist vis-à-vis that problem.

But before concluding, we must notice a change of philosophical perspective in the Radical Anti-realist discourse. Indeed the reference to the 'Aristotelian ontology of infinities' in Marion's paper quoted above, cannot be read as a strict finitist, because it appeals to the potential infinite. So either one abandons the radical spirit of this sort of anti-realism, or this last move betrays what I would call a 'philosophical unstability'. Last but not least, this argument shows the benefits of a philosophical analysis starting from well defined philosophical systems: from the point

---

[15] Marion in [8, p. 423].

[16] See [7].

of view of Vuillemin's classification, an 'Aristotelian ontology of infinities' has a 'dogmatic' connotation, and if this 'dogmatism' can be avoided, it remains that this expression is not compatible with the strict finitism.

17.3 Conclusion: Laziness or Heroism?

We have just seen that, unfortunately, neither point (2) nor point (3) led convincingly to a logic that could be the logic of feasible mathematical problems, if we expect, of course, that such a logic has the handy property of decidability: for every finite structure C one should be able to decide if C definable in the logic of P or not. But, unfortunately, a philosopher can always find a way out and avoid difficult problems, instead of finding solutions. The lazy attitude for the Radical Anti-realist would consist in giving up the project of defining *the* logical system for her philosophy. The Radical Anti-realist would be then the troublemaker for all philosophers of logic, reminding them obstinately of the platitude that the only real proofs that one gets are always *finite proofs*. The Radical Anti-realist would degenerate into a sort of pragmatism which forgets the philosophical search for the kernel of the logic of knowledge, and considers mathematical logic as various tools for various tasks, with more or less efficient algorithms.

To such laziness, the 'heroic' attitude for Radical Anti-realist would be to believe in finding the logic of the complexity class P, and thus to bet that $\mathrm{P} = \mathrm{NP}$ can be proved by the same fact, contrarily to the the orthodox conjecture which is presently that $\mathrm{P} \neq \mathrm{NP}$. If that is the direction taken by Marion and Dubucs, one must recognize that their Radical Anti-realism has the double merit of not being dependent of an orthodoxy, and of stressing the philosophical importance of the big scientific question of complexity theory. It is also true that if the logic of feasibility would be discovered (or, if one prefers, would be invented), it could finally help us to abandon Platonism more frankly than good old intuitionism did.

But independently of replies to the points (2) and (3), a reply to the point (1), i.e. to the 'Revolutionary Argument' itself, must be given to close this paper. The 'Revolutionary Arguments' seems to forget that *the first real rupture* from Platonism (and generally from Dogmatism, to speak like Vuillemin) has been done by intuitionism. There is something unfair in the criticism that the Radical Anti-realist makes to intuitionism. As Girard [6, p. 4] has rightly noticed, "there is no constructive logic beyond intuitionistic logic"; that means clearly that the demarcating line between Platonism in mathematics and constructivist mathematics has been drawn by the intuitionistic logic. Of course, one can hope a sharper distinction like the Radical Anti-realist requires. But to take again the complexity question, one knows that the intuitionistic logic is PSPACE-complete while the classical logic is NP-complete.[17] It means that the former makes the decision problem harder. Consequently, provided

[17] [11], quoted in Tennant [12, p. 5].

the feasibility requirement is not given up, the Radical Anti-realist meets the problem of defining a logic which does not make the decision problem still harder than in intuitionistic logic.

Acknowledgements Many thanks to the editors of this volume for their encouragement to publish this paper. I am especially thankful to Mathieu Marion for having revised my English.

References

1. Bell, J. L., D. DeVidi, and G. Solomon. 2001. *Logical Options: An Introduction to Classical and Alternative Logics*. Peterborough, ON: Broadview Press.
2. Bridges, D. 1999. “Can Constructive Mathematics Be Applied in Physics?” *Journal of Philosophical Logic* 28(5):439–53.
3. D’Agostino, M., and D. Gabbay, and K. Broda. 1999. “Tableaux Methods for Substructural Logics.” In *Handbook of Tableaux Methods*, edited by M. D’Agostino, D. M. Gabbay, R. Hänle, and J. Possega, 397–467. Dordrecht: Kluwer.
4. Dubucs, J. 2002. “Feasibility in Logic.” *Synthese* 132:213–37.
5. Dubucs, J., and M. Marion. 2003. “Radical Anti-realism and Substructural Logics.” In *Philosophical Dimensions of Science. Selected Contributed Papers from the 11th International Congress of Logic, Methodology, and the Philosophy of Science, Krakow, 1999*, edited by A. Rojszczak et al., 235–49. Dordrecht: Kluwer.
6. Girard, J.-Y. 1987. “Linear Logic.” *Theoretical Computer Science* 50(1):1–102.
7. Immerman, N. 2008. “Computability and Complexity.” *Stanford Encyclopedia of Philosophy*, 2004–2008. http://plato.stanford.edu/entries/computability/.
8. Marion, M. 2009. “Radical Anti-realism, Wittgenstein and the Length of Proofs.” *Synthese* 171:419–32.
9. D’Agostino, M., D. M. Gabbay, R. Hänle, and J. Possega, eds. 1999. *Handbook of Tableaux Methods*. Dordrecht: Kluwer.
10. Russell, B. 1940. *An Inquiry into Meaning and Truth*. London: Allen & Unwin.
11. Statman, R. 1979. “Intuitionistic Propositional Logic is Polynomial Space Complete.” *Theoretical Computer Science* 9(1):67–72.
12. Tennant, N. 1992. *Autologic*. Edinburgh Information and Technology Series. Edinburgh: Edinburgh University Press.
13. Tennant, N. 1997. *The Taming of the True*. Oxford: Oxford University Press.
14. Tennant, N. 2006. “Logic, Mathematics and the Natural Sciences.” In *Handbook of the Philosophy of Science*, vol. 5, Philosophy of Logic, edited by D. Jaquette, 1149–66. Amsterdam: Elsevier BV.
15. van Bendegem, J.-P. 2002. “Finitism in Geometry.” *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/.
16. Vuillemin, J. 1984. *Nécessité ou contingence, l’aporie de Diodore et les systèmes philosophiques*. Paris: Minuit.
17. Vuillemin, J. 1986. *What Are Philosophical Systems?* Cambridge, MA: Cambridge University Press.
18. Vuillemin, J. 1996. *Necessity or Contingency—The Master Argument*. Number Lecture Notes 56. Stanford, CA: CSLI Publications.
19. Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul Ltd. trad. fr. Granger, Gallimard, Paris, 1993.

Chapter 18
Two Diamonds Are More Than One

Transitivity and the Factivity of Feasible Knowability

Elia Zardini

18.1 Introduction and Overview

Semantic anti-realism (henceforth simply 'anti-realism')[1] may neutrally be characterized as the doctrine that *there is a conceptual connection between truth and our recognition of it*. As qualifiedly applied to a particular discourse D, anti-realism is the doctrine that there is a conceptual connection between the truth of the sentences belonging to D and our recognition of it.[2] This conceptual connection is quite naturally supposed to be captured by the formulation of an *epistemic constraint* on the notion of truth operating over a certain discourse:

(EC$^-$) Necessarily, 'P' is true only if it is feasibly knowable that P,[3]

where 'necessarily' expresses metaphysical necessity.

Two features of the modal epistemic operator 'it is feasibly knowable that', as used by the anti-realist in stating her thesis, are worth remarking upon right at the outset. Firstly, the operator is intended to be *factive*: that is, its being feasibly knowable that P entails that P (but see [15, pp. 230–37] for a recent disagreeing voice). In the presence of the *disquotational* schema:

[1] Throughout, I use simple quotation marks for standard (and shudder) quotation, corner quotes and display for autonymous quasi-quotation (where non-metalinguistic quantification into such environment has to be understood substitutionally). Formal vocabulary refers to itself.

[2] Call the unqualified version of anti-realism, quantifying over every sentence belonging to *whichever* discourse, 'global anti-realism'. Call a qualified version of anti-realism, quantifying just over every sentence belonging to a *particular* discourse D, $\ulcorner$local anti-realism with respect to $D\urcorner$. The discussion in the paper is intended to be largely insensitive to this distinction.

[3] Throughout, 'P' will be used as a substitutional sentential variable (if free, amounting in effect to a sentence schema). 'φ' will instead be used as a metalinguistic variable over sentences.

E. Zardini (✉)
Department of Philosophy, Northern Institute of Philosophy, University of Aberdeen, Aberdeen, UK
e-mail: elia.zardini@abdn.ac.uk

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1_18,

(D) Necessarily, 'P' is true iff P[4]

(where, again, 'necessarily' expresses metaphysical necessity), (EC^{-}) can therefore be strengthened to:

(EC) Necessarily, 'P' is true iff it is feasibly knowable that P.

Secondly, that the knowability in question is a *feasible* one is typically taken to mean that the relevant metaphysically possible situations which verify a claim of feasible knowability are situations:

(i) which concern epistemic subjects endowed with our actual cognitive powers or, at most, with some *finite* extensions thereof;
(ii) in which the available evidence is at least nomologically constrained by the *present* situation of the world (that is, by the situation at which the claim of feasible knowability is to be evaluated).[5]

(i) ensures that (EC) is not made trivial e.g. for arithmetical discourse by the metaphysical possibility of a being endowed with an *infinite* extension of our actual cognitive powers, being which would thus be in a position to survey 'at a glance', as it were, the whole number series in its infinity and so in a position to decide every sentence concerning it (just as we are in a position to decide every quantifier-free sentence concerning it). (ii) ensures that (EC) requires e.g. nothing less than nomologically available evidence in order for a past-tensed sentence to be true, and so, presumably, *traces in the present* of the past facts the sentence is about.[6]

Moreover, I henceforth assume that the modality expressed by 'it is feasibly knowable that' results from the composition of the epistemic modality of knowledge with a *sui generis* alethic modality, and call the latter '*feasible possibility*' (and its dual '*feasible necessity*'). Feasible knowability is then just feasible possibility of knowledge.[7] Since inspection of (i) and (ii) reveals that the possible situations relevant for a claim of feasible possibility result in effect from a *restriction* on the collection of all metaphysically possible situations, the collection of all feasible possibilities will properly be included in the collection of all metaphysical possibilities.

---

[4] Possible restrictions on the schema in order to deal with context sensitivity and the semantic paradoxes will be ignored here, as not relevant to the issues at hand (see [21] and [22] for respective discussions of these two problems).

[5] I leave it intentionally vague what exactly the strength of such constraint is. Especially in deterministic worlds, the constraint should not presumably be so strong as to entail that the situation in question be *in all its features* nomologically possible given the present situation of the world.

[6] I should stress that neither (i) nor (ii) are non-negotiable for a broadly anti-realist perspective. I only report them as conditions typically occurring in anti-realist discussions (as can be variously found e.g. in [5]) and they will actually play almost no role in the rest of my discussion, which focusses instead on the factivity feature. Thanks to an anonymous referee for raising this point.

[7] 'Knowledge' and its relatives are in turn henceforth understood as implicitly existentially generalizing over epistemic subjects and times: knowledge by *someone* at *some time*.

The rest of the paper is organized as follows. Section 18.2 expounds the problem posed by a simple well-known argument to a certain version of anti-realism, and considers a different version of it. Section 18.3 presents a novel challenge posed to the modified version, challenge which calls for a deeper understanding of the semantics of the operator ‘it is feasibly possible that’. Section 18.4 undertakes this task, showing how the attendant accessibility relation for the operator must be relativized to particular facts, and how this relativization can lead to peculiar failures of transitivity. In the lights of this result, Section 18.5 considers (non-transitive) epistemic possibility as candidate for playing the role of feasible possibility, and finds it wanting. Section 18.6 draws the conclusions which follow for anti-realism and, more generally, for our understanding of the modality expressed by the operator ‘it is feasibly possible that’.

18.2 The Paradox of Knowability and the Restriction Strategy

In general, *unrestricted* anti-realism is the doctrine that, for whichever schema (such as (EC)) is to constrain the truth of sentences belonging to a discourse, *every* instance of it holds.[8] Under some natural assumptions, *(EC)*-unrestricted anti-realism is refuted by the following simple argument, originally published in [7, pp. 138–39][9] and long known as ‘*the paradox of knowability*’:

P_1	(P_1)	$\mathcal{K}(\varphi \wedge \neg\mathcal{K}\varphi)$	A
P_1	(P_2)	$\mathcal{K}\varphi$	P_1 $\mathbf{IKT}_{\mathcal{K}}$
P_1	(P_3)	$\mathcal{K}\neg\mathcal{K}\varphi$	P_1 $\mathbf{IKT}_{\mathcal{K}}$
P_1	(P_4)	$\neg\mathcal{K}\varphi$	P_3 $\mathbf{IKT}_{\mathcal{K}}$
P_1	(P_5)	$\mathcal{K}\varphi \wedge \neg\mathcal{K}\varphi$	P_2,P_4 **IP**
	(P_6)	$\neg\mathcal{K}(\varphi \wedge \neg\mathcal{K}\varphi)$	P_1,P_5 **IP**
	(P_7)	$\Box\neg\mathcal{K}(\varphi \wedge \neg\mathcal{K}\varphi)$	P_6 $\mathbf{IKT}_{\Box}$,

where $\Box$ expresses metaphysical necessity, $\mathcal{K}\varphi$ formalizes ⌜Someone knows at some time that φ⌝, **IP** is intuitionist propositional logic, $\mathbf{IKT}_{\Box}$ an intuitionistically acceptable version of a **KT** logic for $\Box$[10] and, analogously, $\mathbf{IKT}_{\mathcal{K}}$ an intuitionistically acceptable version of a **KT** logic for $\mathcal{K}$. Given that metaphysical necessity (of lack of knowledge) entails feasible necessity (of lack of knowledge), metaphysical necessity of lack of knowledge intuitionistically entails the negation of feasible

[8] Note that this distinction between unrestricted and restricted anti-realism is orthogonal to the distinction between global and local anti-realism introduced in fn. 2.

[9] But most likely due to Alonzo Church, see [11, pp. 34–37].

[10] Unfortunately, there is no unique such version. See [1, 4] for some plausible proposals.

possibility of knowledge.[11] Thus, (EC)-unrestricted anti-realism in conjunction with (P_7) entails that, for every φ, $\varphi \wedge \neg\mathcal{K}\varphi$ is not true. Taking φ to be 'P', disquoting with (D) and generalizing in sentence position, we obtain from this the result that there are no unknown truths.[12] Classically, this result is logically equivalent to the statement that every truth is known. If you accept classical logic and think that it is not the case that every truth is known, you had better reject (EC)-unrestricted anti-realism. If you accept intuitionist logic and think that there are unknown truths, you too had better reject (EC)-unrestricted anti-realism (and, of course, you had better reject it even if you only accept that it is not the case that there are no unknown truths).

Anti-realism is a deep and difficult philosophical doctrine. It would be very strange, to say the least, if it were once and for all refuted by such a simple piece of reasoning as (P_1)–(P_7) in conjunction with (D) and the *epistemic-modesty* principle:

(EM) For some P, P and it is not known that P.

The correct reaction to the paradox of knowability has therefore seemed to be to many commentators that of concluding that what the paradox shows is, at most, that a theorist who endorses (EC)-unrestricted anti-realism has unnecessarily overstated the idea characteristic of anti-realism, and thereby committed only, as it were, a tactical mistake. After all, it should have come as no surprise, given the level of *complexity* induced on a primitive language by the introduction of logical operators, that their interactions with the atomic sentences of the language and with an epistemic operator of some kind or other ends up yielding logically or conceptually unknowable truths. For, in general, an epistemic subject s's coming to be in some epistemic relation E (not necessarily the one of knowledge)[13] to a content (c_0) may very well logically or conceptually entail that E does not hold between s and the content (c_1) that E does not hold between s and c_0. Thus, assuming that E's holding between s and a conjunction logically or conceptually entails E's holding between s and each conjunct, E cannot, on purely logical or conceptual grounds, hold between s and the conjunction of c_0 and c_1, even if such a conjunction may very well be true.

These considerations may seem to point towards a very weak *restriction* on (EC) (along the lines of that favoured in [12, pp. 245–79]; [13, 16, pp. 11–21] and defended—successfully, in my opinion—in [14] against the stricter restriction to

[11] To see that this modal move from the feasible necessity of a negation to the negation of a feasible possibility should count as intuitionistically valid, think of operators of feasible modality as first-order quantifiers over the range of feasibly possible worlds. Then the *modal* move in question amounts to the intuitionistically kosher *first-order* move from $\forall\xi\neg\varphi$ to $\neg\exists\xi\varphi$.

[12] Henceforth, I mostly use the word 'truth' merely as a handy short for the clumsy substitutional locution 'thing which is the case'. Thus, for example, 'There are no unknown truths' is understood to mean 'For no P, P and it is unknown that P'.

[13] For example, as Mackie [9, p. 91] first noted, it seems just as conceptually impossible to be justified in believing that [P and no one is ever justified in believing that P] as it is to know that [P and no one ever knows that P].

atomic sentences favoured in [6])[14]: namely, the restriction that, if $\ulcorner P \urcorner$ is to be a legitimate *substituens* in the (EC)-schema, it be metaphysically possible to know that P (see the sustained exchange [13, 16, 19, 20] for further discussion). The gist of anti-realism may seem to be that, if something is true *and it is metaphysically possible to know it*, then we as we actually are (or some finite extension of us as we actually are) can know it in our present evidential situation. Of course, the challenge remains for such and similar restriction strategies to locate, in the arguments given for anti-realism, the (implicit and) principled restriction which makes it the case that these arguments, if successful at all, only establish the favoured restricted version of (EC) rather than unrestricted (EC) itself (cf. [19, pp. 112–13]). This challenge is certainly very pressing and I myself would actually favour a different approach on behalf of the anti-realist (as detailed in [23]). Be that as it may, in the rest of this paper I would like to address instead another recent objection raised against restriction strategies.

18.3 A New Threat of Collapse of Feasible Knowability on Actual Knowledge

Berit Brogaard and Joe Salerno have recently argued that, roughly, even versions of anti-realism relying on *any* kind of restriction on (EC) are threatened with inconsistency with (EM) interpreted as quantifying over the same range to which the restricted version of (EC) is still supposed to apply.[15] No progress would then have been made from the original situation of inconsistency between (EC)-unrestricted anti-realism and (EM) revealed by the paradox of knowability (see [2, 3, pp. 17–18]). Such a conclusion is, I submit, *incredible*: how could it ever be that our notion of feasible knowability is per se so strong as to rule out that *any* restricted version of anti-realism applies—even only de facto—to a class of sentences for which (EM) also holds?[16] In other words, how could it ever be that our notion of a sentence's being feasibly knowable inevitably collapses (classically) on the notion of that sentence's being known or at least (intuitionistically) on the notion of that sentence's not being unknown?

Any such case is clearly bound to overkill: if the notion of a feasibly knowable sentence which is not known is nothing but a *contradictio in adiecto*, then *no one*—i.e. neither the anti-realist *nor the realist*—can hold that it is not the case that every

[14] For the record, [15,16, pp. 21–23] proposes a related but interestingly different restriction.

[15] Strictly speaking, that requires a qualification, since some not completely trivial assumptions are made about what the restriction is (in the argument of Section 18.4, it is assumed that $\mathcal{K}\varphi$ is subject to (EC) if φ is; in the argument of Section 18.5, it is assumed that $\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi$ and $\neg\varphi$ are subject to (EC) if φ is). These assumptions are however so minimal that, in what follows, I will mostly leave the qualification implicit.

[16] Think especially of the limit-case in which the class of sentences is the singleton of an arbitrarily chosen true but unknown sentence.

unknown truth is not feasibly knowable and, a fortiori, that there are feasibly knowable truths which are not known (cf. [10, p. 70]). But that is highly counterintuitive, to say the least. For instance, it seems very plausible that there is a metaphysically possible world w_0 where, for some P, at an earlier time t_0 we have not yet known that P while at a later time t_1 we discover that P: knowledge develops in time. However, assuming that the laws of w_0 are indeterministic, it also seems very plausible that there is a metaphysically possible world w_1 nomologically (accessible from w_0 and) accessing w_0 where every relevant epistemic subject is killed by, say, a natural catastrophe soon after t_0 and before t_1, and where no relevant epistemic subject is to come into existence in the further history of the world. Clearly, w_0 at t_1 constitutes a metaphysically possible situation satisfying the conditions (i) and (ii) for verifying a claim of feasible knowability at the situation constituted by w_1 at t_0. Thus, w_1 would be a world where it is feasibly knowable that P even though it is not known by anyone at any time that P. What Brogaard and Salerno's case, if sound, would show is that, surprisingly enough, we can rule out a priori that such a w_1 is a genuine possibility!

Indeed, Brogaard and Salerno's case, if sound, would be a devastating case against what we may dub 'optimistic epistemic modesty' (the claim, usually endorsed both by realists and anti-realists, that there are unknown truths which are nevertheless feasibly knowable), underwriting either a form of 'epistemic despair' (for every truth you think you have more reason to believe not to be known than to be feasibly knowable: each such truth would then not be feasibly knowable either) or a form of 'epistemic arrogance' (for every truth you think you have more reason to believe to be feasibly knowable than not to be known: each such truth would then not be unknown either, and therefore, classically, would indeed be known).[17]

It is therefore very likely that Brogaard and Salerno's case is flawed. Their overarching case divides in three different independent arguments, all of which are aimed at establishing, in different ways, the result that any kind of restriction on (EC) is inconsistent with (EM) interpreted as quantifying over the same range to which the restricted version of (EC) is still supposed to apply. Maybe sensing the devastating force that their overarching case, if sound, would possess, they cautiously conclude that '[. . .] the restriction strategies proposed thus far are insufficient to treat the real problem. The paradoxes presented herein turn on the basic logic of $\Diamond$ and the ways in which $\Diamond$ operates on epistemic statements. If a restriction strategy can be vindicated, this will be known only after we have formally analysed the anti-realist's notion of possibility' [2, p. 149]. The rest of this paper is intended to be a first step in that direction. Assuming a broadly possible-worlds semantics for the feasible-possibility operator, I will set forth, in the lights of the informal characterization given at the outset of the paper, some features that a good formal semantics for the operator can be expected to have. This will already be sufficient to show the unsoundness of two

[17] Of course, relative to the class of sentences for which they believe that (EC) holds good, the mandated attitude for anti-realists would be epistemic arrogance.

of the arguments offered by Brogaard and Salerno and, more importantly, will provide the beginnings of an appreciation of the distinctive modality expressed by the operator ‘it is feasibly possible that’.

18.4 Transitivity, Factivity, and the Relativity of Accessibility

In the following, $\blacklozenge$ formalizes the feasible-possibility operator (and hence $\blacklozenge\mathcal{K}$ formalizes the feasible-knowability operator), while $\blacksquare$ formalizes its dual. The first argument assumes $\mathcal{K}\varphi$ to be subject to (EC) if φ is, and runs as follows (adapted from [2, pp. 144–45])[18]:

T_1	(T_1)	$\varphi \wedge \neg\mathcal{K}\varphi$	A
T_2	(T_2)	$\blacksquare(\mathcal{K}\varphi \equiv \blacklozenge\mathcal{K}\mathcal{K}\varphi)$	(EC), (D) **IKT**$_\blacksquare$
T_2	(T_3)	$\mathcal{K}\varphi \equiv \blacklozenge\mathcal{K}\mathcal{K}\varphi$	T_2 **IKT**$_\blacksquare$
T_4	(T_4)	$\Box(\varphi \equiv \blacklozenge\mathcal{K}\varphi)$	(EC), (D) **IKT**$_\Box$
T_4	(T_5)	$\varphi \equiv \blacklozenge\mathcal{K}\varphi$	T_4 **IKT**$_\Box$
T_1	(T_6)	$\neg\mathcal{K}\varphi$	T_1 **IP**
T_1,T_2	(T_7)	$\neg\blacklozenge\mathcal{K}\mathcal{K}\varphi$	T_3,T_6 **IP**
T_1	(T_8)	φ	T_1 **IP**
T_1,T_4	(T_9)	$\blacklozenge\mathcal{K}\varphi$	T_5,T_8 **IP**
T_2	(T_{10})	$\blacklozenge\mathcal{K}\varphi \equiv \blacklozenge\blacklozenge\mathcal{K}\mathcal{K}\varphi$	T_2 **IKT**$_\blacksquare$
T_1,T_2,T_4	(T_{11})	$\blacklozenge\blacklozenge\mathcal{K}\mathcal{K}\varphi$	T_{10},T_9 **IP**
T_1,T_2,T_4	(T_{12})	$\blacklozenge\mathcal{K}\mathcal{K}\varphi$	T_{11} **IKT4**$_\blacksquare$
T_1,T_2,T_4	(T_{13})	$\blacklozenge\mathcal{K}\mathcal{K}\varphi \wedge \neg\blacklozenge\mathcal{K}\mathcal{K}\varphi$	T_{12},T_7 **IP**
T_2,T_4	(T_{14})	$\neg(\varphi \wedge \neg\mathcal{K}\varphi)$	T_1,T_{13} **IP**,

where φ is an arbitrary sentence falling under the scope of the restricted version of (EC), **IKT**$_\blacksquare$ an intuitionistically acceptable version of a **KT** logic for $\blacksquare$ and **IKT4**$_\blacksquare$ an intuitionistically acceptable version of a **KT4** logic for $\blacksquare$. (T_{14}) is inconsistent with (EM)’s holding over the range to which the restricted version of (EC) is still supposed to apply.

The operator ‘necessarily’ occurring in both (EC) and (D) expresses metaphysical rather than feasible necessity. Nevertheless, (EC) and (D) jointly justify (T_2) since, as already noted, metaphysical necessities are (properly) included in feasible necessities. Furthermore, the closure step at line (T_{10}) seems to be justified as well.[19] The problem with (T_1)–(T_{14}) may then reasonably be thought to lie in the step from (T_{11}) to (T_{12}), which makes use of the $\blacklozenge$-version of the characteristic **KT4** axiom ($\blacksquare\varphi \supset \blacksquare\blacksquare\varphi$), the *transitivity* of the relevant accessibility relation being the necessary and sufficient condition for its semantic justification. The question then is whether the correct semantics for $\blacklozenge$ should employ a transitive accessibility

[18] I should like to stress that the kernel of the argument goes actually back to [17, p. 67].

[19] See fn. 26 for a more precise statement concerning (T_{10}).

relation. Brogaard and Salerno too seem to identify this as the crux of (T_1)–(T_{14}): 'One might object that we simply need to treat $\Diamond$ as non-transitive. But this has not yet been argued for. And it would not be very interesting simply to suppose the non-transitivity of $\Diamond$, having no reason other than the threat of the revised Fitch paradox to motivate the supposition. Pending further discussion, the supposition of non-transitivity is ad hoc' [2, p. 145].

I accept Brogaard and Salerno's request for a *principled* rejection of transitivity. However, a moment's reflection does suffice to show that the intuitive notion of feasible knowability, as spelled out at the outset of this paper, determines that the semantics of $\blacklozenge$ cannot allow for an unrestrictedly transitive accessibility relation. We start by observing that the *factivity* constraint determines that the only worlds[20] which are relevant in the evaluation of a claim of feasible knowability (that is, the only worlds which are *feasibly accessible*) are those worlds in which we hold *everything constant* (with respect to the world at which the claim of feasible knowability is to be evaluated) *except for the relevant epistemic facts* (and their presuppositions and consequences).[21] This is meant to ensure the factivity of the complex modal epistemic operator $\blacklozenge\mathcal{K}$, and thereby to ensure its adequacy as a formalization of the operator 'it is feasibly knowable that'. For what may be regarded as merely possible and non-actual[22] in a claim that it is feasibly knowable that P is *not* the fact that P (which is, on the contrary, assumed actually to obtain), but simply the fact that *it is known that* P. What a claim that it is feasibly knowable that P entails is that it is actually the case that P and that, although it may not actually be known that P, the present state of the actual world determines an evidential situation such that a being endowed with our actual cognitive powers—or with some finite extensions thereof—can—under the constraints of the actual laws of nature—come to know that P.

With such restrictions on feasible accessibility in place, principled counterexamples to its unrestricted transitivity are forthcoming. Consider, for instance, worlds w_0, w_1 and w_2. In w_0, there are 1,963 houses in Nancy, but it is not known that there are 1,963 houses in Nancy. w_1 holds constant every fact of w_0 with the exception that in w_1 it is known that there are 1,963 houses in Nancy. w_2 holds constant every fact of w_1 possibly with the exception that in w_2 the relevant epistemic subjects reflect on their cognitive achievements and thereby come to know that it is known

[20] Henceforth, I will set aside complications arising from time.

[21] This distinction between the relevant epistemic facts with their presuppositions and consequences on the one side and the rest of all other facts on the other side presupposes a mild form of non-holism. Although such issues clearly go beyond the scope of this paper and I won't try to defend this claim here, it is arguable that it is under that presupposition that the concept of feasible knowability finds its usefulness in ordinary and philosophical thinking. Thanks to Jonathan Lowe, Sven Rosenkranz and an anonymous referee for discussions on this point.

[22] Note that, as has just happened in the text, I sometimes use 'actually' and its relatives not just to refer to the actual world, but, more generally, to refer back to whichever world is the world at which a certain claim is to be evaluated.

that there are 1,963 houses in Nancy. As the counterexample we are constructing is neutral with respect to the validity of the so-called '$\mathcal{K}\mathcal{K}$-*thesis*':

(KK) If it is known that P, then it is known that it is known that P,

w_2 is not assumed to be distinct from w_1 (if the reader thinks that (KK) does not hold, then she is free to assume that w_2 is distinct from w_1; if, on the contrary, she thinks that (KK) does hold, then she is free to assume that w_2 is identical with w_1). Assume also that w_0, w_1 and w_2 all satisfy (i) and (ii) with respect to one another. Then, we may take it, *relative to* there being 1,963 houses in Nancy, w_1 is feasibly accessible from w_0, and, *relative to* its being known that there are 1,963 houses in Nancy, w_2 is feasibly accessible from w_1. Assigning 'It is known that there are 1,963 houses in Nancy' to 'φ', $\mathcal{K}\varphi$ is true at w_2, from which it follows that $\blacklozenge\mathcal{K}\varphi$ is true at w_1 and that $\blacklozenge\blacklozenge\mathcal{K}\varphi$ is true at w_0 (see fn. 25 for a justification of this very last claim). Does it also follow that $\blacklozenge\mathcal{K}\varphi$ is true at w_0? Emphatically no, because that would in effect mean that, relative to its being known that there are 1,963 houses in Nancy, w_2 is feasibly accessible from w_0, which it isn't, since, in w_0, it is not known that there are 1,963 houses in Nancy while, in w_2, it is, and hence that fact—which, relative to itself, is clearly not one of the relevant epistemic facts exempt from the requirement of being held constant—is not held constant from w_0 to w_2.

The lesson implicit in this kind of example can be elaborated as follows. A world w_j is feasibly accessible or not from a world w_i only *relative to* a particular fact candidate for being feasibly knowable at w_i. In particular, the factivity constraint determines that w_j is feasibly accessible from w_i relative to a particular fact *only if* that fact obtains both at w_i and w_j (call the facts which meet this condition $\ulcorner w_i$-w_j-constant$\urcorner$). $\blacklozenge$ in $\blacklozenge\mathcal{K}\varphi$ can then be interpreted as selecting as relevant feasible-accessibility relation the specific one relative to the fact expressed by φ. Thus, for instance, in our counterexample w_2 is feasibly accessible from w_0 relative to the fact that there 1,963 houses in Nancy, as this fact is indeed w_0-w_2-constant, but, as we have just noted, it is not feasibly accessible from w_0 relative to the fact that it is known that there 1,963 houses in Nancy, as this latter fact is not w_0-w_2-constant. Consequently, whilst w_2 can count as a witness for the feasible possibility, in w_0, that it is known that there 1,963 houses in Nancy, it cannot count as a witness for the feasible possibility, in w_0, that it is known that it is known that there 1,963 houses in Nancy. Indeed, since for no w_j is the fact that it is known that there 1,963 houses in Nancy w_0-w_j-constant (for that fact does not obtain at w_0 in the first place!), no world can count as such a witness and so, at w_0, it is not the case that it is feasibly possible to know that it is known that there 1,963 houses in Nancy, as desired.

The facts which are not w_i-w_j-constant will include epistemic facts of the relevant order n, which obviously have to be allowed to change from not obtaining at w_i to obtaining at w_j if the notion of feasible possibility (and, in particular, of feasible knowability) is to have any interest at all and not to collapse on that of actuality (and, in particular, of actual knowledge). However, the very same nth-order epistemic facts which so vary from w_i to w_j may very well, with respect to a world w_k which meets the other conditions for being feasibly accessible from w_j relative to any nth-order epistemic fact, be w_j-w_k-constant.[23] Importantly, any such w_j-w_k-constant fact will not be w_i-w_k-constant. Thus, for instance, in our counterexample the first-order epistemic fact that it is known that there 1,963 houses in Nancy is not w_0-w_1-constant, and yet is w_1-w_2-constant.[24] Importantly, such w_1-w_2-constant fact is not w_0-w_2-constant. It is then clear that, under such circumstances, one cannot conclude without further qualification, from the holding of these two *relativized* feasible-accessibility relations between w_j and w_i and between w_k and w_j, that w_k is also feasibly accessible from w_i relative to *every* fact whatsoever. For, whereas w_k may very well be assumed to be, in turn, feasibly accessible from w_i *relative to* every w_i-w_j-constant fact, it cannot be assumed to be feasibly accessible from w_i *relative to* every w_j-w_k-constant fact, for some such fact (namely, the holding of some epistemic relation to some w_i-w_j-constant fact) may not be w_i-w_k-constant.

Crucially, this peculiar kind of failure of transitivity suffices to determine failures of the characteristic **KT4** axiom. For collapse at a world w_i of $\blacklozenge\blacklozenge\varphi$ on $\blacklozenge\varphi$ (where $\blacklozenge\blacklozenge\varphi$ is true at w_i because $\blacklozenge\varphi$ is true at w_j[25] and this in turn because φ is true at w_k) is ensured to be truth preserving only if it is not the case that φ describes the holding of an epistemic relation to a fact which, while w_j-w_k-constant, is not w_i-w_j-constant. For, if such requirement were not met, that fact would not be w_i-w_k-constant either, and so w_k would not be feasibly accessible from w_i *relative to it*, and so w_k could not count as a witness to the truth of $\blacklozenge\varphi$ at w_i. Thus, under this requirement, one cannot, in the context of (T_1)–(T_{14}), collapse the sentence $\blacklozenge\blacklozenge\mathcal{K}\mathcal{K}\varphi$ on $\blacklozenge\mathcal{K}\mathcal{K}\varphi$ (as at line (T_{12})), for, in the context of (T_1)–(T_{14}), $\mathcal{K}\mathcal{K}\varphi$ describes exactly the holding of an epistemic relation to a fact (expressed by $\mathcal{K}\varphi$) which, while w_j-w_k-constant, is not w_i-w_j-constant.[26]

---

[23] Notice though that the epistemic facts which do change from not obtaining at w_j to obtaining at w_k may still include the $n+1$th-order epistemic facts constituted by the holding of epistemic relations to the nth-order epistemic facts in turn constituted by the holding of those epistemic relations which have varied from w_i to w_j but which are held constant from w_j to w_k.

[24] Notice though that what does change from not obtaining at w_1 to obtaining at w_2 may still include, if (KK) does not hold, the second-order epistemic fact that it is known that it is known that there 1,963 houses in Nancy.

[25] In our counterexample, we have assumed that w_1 can count as a witness for the feasible possibility, in w_0, not only that it is known there are 1,963 houses in Nancy, but also that it is feasibly knowable that it is known that there are 1,963 houses in Nancy. The assumption is warranted by the consideration that those two facts (i.e. the fact that it is known there are 1,963 houses in Nancy and the fact that it is feasibly knowable that it is known there are 1,963 houses in Nancy) would seem to belong to the same order, since they would seem to concern the same level of actual knowledge (namely, the first one) and by the consideration that, if two facts belong to the same order, a world is feasibly accessible from a world relative to one fact iff it is so accessible relative to the other fact.

[26] These considerations concerning the relativity to particular facts of the feasible-accessibility relation can be developed in a formal framework whose main characteristic is the use of *multiple* feasible-accessibility relations. The facts to which feasible accessibility is relative can in effect—with some simplification—be merged together in the following sets: the set of non-epistemic facts, the set of first-order epistemic facts, the set of second-order epistemic facts etc. It is then natural to mirror this hierarchy by classifying formulae according to their degree of complexity with respect

18.5 Epistemic Possibility of Knowledge and Feasible Knowability

I would like to close by considering a proposal made by Brogaard and Salerno for an identification of the modality expressed by $\blacklozenge$ which satisfies the constraints on it unveiled by our foregoing discussion. Since they (correctly, as we have seen in our discussion of (T_1)–(T_{14})) think that a restricted version of (EC) entails, under (EM), that the accessibility relation featuring in the semantics of $\blacklozenge$ is non-transitive, and (incorrectly, as I don't have space to expand on here) think that a restricted version of (EC) is also incompatible, under (EM), with (KK), Brogaard and Salerno eventually cast about for an interpretation of $\blacklozenge$ which, under the rejection of (KK), verifies the non-transitivity of $\blacklozenge$. They rightly note that *epistemic possibility*, defined as '*consistency*' with what is known,[27] behaves non-transitively under the rejection of (KK).

Indeed, it is easily shown that, given the above definition of epistemic possibility, failure of (KK) is both necessary and sufficient for the accessibility relation featuring in the semantics of the epistemic-possibility operator (that is, *epistemic accessibility*) to be non-transitive. As for necessity, observe that, if (KK) holds, then every world w_0 at which $\mathcal{K}\varphi$ is true is a world at which $\mathcal{K}\mathcal{K}\varphi$ (and $\mathcal{K}\mathcal{K}\mathcal{K}\varphi$, and $\mathcal{K}\mathcal{K}\mathcal{K}\mathcal{K}\varphi\ldots$) is true, so that every world w_1 epistemically accessible from w_0 is a world at which (both φ and) $\mathcal{K}\varphi$ (and $\mathcal{K}\mathcal{K}\varphi$, and $\mathcal{K}\mathcal{K}\mathcal{K}\varphi\ldots$) are true, so that every world w_2 epistemically accessible from w_1 is a world at which φ is true (because of the truth at w_1 of $\mathcal{K}\varphi$), and $\mathcal{K}\varphi$ is true (because of the truth at w_1 of $\mathcal{K}\mathcal{K}\varphi$), and $\mathcal{K}\mathcal{K}\varphi$ is true (because of the truth at w_1 of $\mathcal{K}\mathcal{K}\mathcal{K}\varphi$)… Therefore, the necessary and sufficient condition for the epistemic accessibility of w_1 from w_0 (truth at w_1 of what is known at w_0) is satisfied by every world w_2 which satisfies the necessary and sufficient condition for being epistemically accessible from w_1. It then follows that, if w_1 is epistemically accessible from w_0 and w_2 from w_1, so is w_2 from w_0—in other words, if (KK) holds, epistemic accessibility is transitive. Contraposing, if

to occurrences of $\mathcal{K}$, and to assign then a specific feasible-accessibility relation to each degree of complexity. The resulting logics for the feasible-possibility operator are investigated in detail in [24]. Here, I should like to mention that the logics are sufficient to invalidate also the second argument considered by Brogaard and Salerno, which I don't have space to address in this paper. I should also report that, even though the logics are non-normal in the sense that the **K** axiom ($\blacksquare(\varphi \supset \psi) \supset (\blacksquare\varphi \supset \blacksquare\psi)$) does not hold unrestrictedly, they still validate the particular instance used at line (T_{10}). I should finally add that, while transitivity fails in the peculiar sense explained in the text (which suffices to determine failures of the characteristic **KT4** axiom), this is compatible with each of the specific feasible-accessibility relations still being transitive.

[27] Given our previous stipulations about 'knowledge' and its relatives in fn. 7, the notion of epistemic possibility defined in the text is neither *group*- nor *time*-relative. Group-relativity is certainly a crucial feature of the notion expressed by ordinary uses of 'might' and its like, but it is inappropriate when trying to analyze a notion that is not group-relative, such as the notion of feasible possibility studied in this paper. That notion may however be time-relative (if condition (ii) is accepted), in which case one should also focus on a time-relative notion of epistemic possibility when discussing the relation between epistemic possibility of knowledge and feasible knowability. The discussion in this section can easily be adapted to the time-relative case.

epistemic accessibility is not transitive, (KK) does not hold.[28] As for sufficiency, observe that, if (KK) does not hold, then the following model invalidates the transitivity of epistemic accessibility (as [2, pp. 146–47, fn. 7] correctly point out): $\mathcal{K}\varphi$ and $\neg\mathcal{K}\mathcal{K}\varphi$ are true at w_0, φ and $\neg\mathcal{K}\varphi$ are true at w_1, $\neg\varphi$ and $\mathcal{K}\neg\varphi$ are true at w_2 (w_1 is epistemically accessible from w_0, and w_2 is epistemically accessible from w_1, but w_2 is not epistemically accessible from w_0).[29]

However, one may very well wonder whether the contribution made by 'it is feasibly knowable that' to the truth conditions of a sentence is the same as the contribution made by 'it is epistemically possible that it is known that'. Given the above definition of epistemic possibility, if it is not known that it is not known that P, then it is epistemically possible that it is known that P. Since its being feasibly knowable that P entails that it is the case that P, its not being known that it is not known that P should then likewise entail that it is the case that P. But that is clearly incorrect: consider any case where it is false that P, but it is not known that it is not known that P just because, say, everyone confidently but mistakenly believes, on the contrary, that it is known that P. In any such case, it would lead to a straightforward contradiction to require that it be the case that P. Suppose, for instance, that everyone confidently but mistakenly believes that whales are fish and that, as a consequence, everyone also believes that it is known that whales are fish. Assuming a *modicum* of rationality, no one will also believe that it is not known that whales are fish, and hence, assuming that knowledge requires belief, no one will know that it is not known that whales are fish. Of course it still doesn't follow that, in such a situation, whales are fish, even though it is epistemically possible to know that they are.[30] Epistemic possibility of knowledge does not sustain factivity.

We then have that the two modal epistemic operators 'it is feasibly knowable that' and 'it is epistemically possible to know that' are very plausibly not even *extensionally* equivalent (and certainly, in any event, not *intensionally* equivalent), for it is very plausibly *actually* the case that some instance of the schema 'It is epistemically possible to know that P' is true whereas the corresponding instance of

---

[28] The gloss in the text 'truth at w_1 of what is known at w_0' clarifies the way in which the rather ambiguous consistency condition for epistemic possibility is to be understood: what is known at a world must be *true* at every world epistemically accessible from it (and not just: what is known at a world must *not* be *false* at every world epistemically accessible from it). Note that the weaker interpretation of the consistency condition would belie failure of (KK) as a necessary condition for the non-transitivity of epistemic accessibility. The following (non-classical) model shows how, under the weaker interpretation of the consistency condition, transitivity fails even under (KK) conditions: $\mathcal{K}\varphi$ and $\mathcal{K}\mathcal{K}\varphi$ are true at w_0, neither $\mathcal{K}\varphi$ nor $\neg\mathcal{K}\varphi$ are true at w_1, $\neg\varphi$ is true at w_2 (w_1 is epistemically accessible from w_0, and w_2 is epistemically accessible from w_1, but w_2 is not epistemically accessible from w_0). Of course, if the semantics is classical, the weaker interpretation of the consistency condition collapses on the stronger. I will ignore such niceties in what follows.

[29] Notice that, in the model described in the text, w_1 and w_2 also invalidate the *symmetry* of epistemic accessibility, and this independently of the question whether (KK) holds or not.

[30] The counterexample needn't rely on the belief requirement for knowledge. For example, it is certainly metaphysically possible that no one has any evidence whatsoever concerning its not being known that whales are fish.

the schema 'It is feasibly knowable that P' is false (and, in any event, it is certainly *metaphysically possible* that some instance of the former schema is true whereas the corresponding instance of the latter schema is false).[31]

The previous counterexample assumed that consistency with what is known (on some interpretation or other) is the necessary *and sufficient* condition for epistemic possibility. Yet, the sufficiency direction of the consistency condition may reasonably be rejected as being too weak, at least in some context (cf. [8, p. 148]). Suppose that a computer collects the results of a complex experiment with subatomic particles which conclusively refute a theory T. Badly misinterpreting the results, I utter ⌜It might be that T is true⌝: my utterance is intuitively false, even though, in that situation, no one knows that T is not true. However, the counterexample can easily be modified to accommodate any plausible stronger condition placed on epistemic possibility, as long as its not being the case that P does not require that it is known that it is not known that P—the vast implausibility of such a requirement can be appreciated by noting that it is equivalent to the characteristic **B** axiom for the knowledge operator ($\varphi \supset \mathcal{K}\neg\mathcal{K}\neg\varphi$), which is notoriously highly problematic (see [18, pp. 166–67, 226–27]).

Epistemic possibility of knowledge is thus not sufficient for feasible knowability. Nor is it necessary. For it may very well be the case that it is feasibly knowable that P but that it is known that it is not known that P (just because, say, it is known that no one will ever bother to check whether P), and therefore, contraposing on the uncontroversial necessity direction of the consistency condition, it is not epistemically possible that it is known that P. Suppose, for instance, that the number of books in my office at noon 28/10/2006 is 92, and that I come to know at noon 28/10/2006 that, alas, the Big Crunch is going to happen in 10 min. Then it is certainly feasibly knowable that the number of books in my office at noon 28/10/2006 is 92 (it would take less than 10 min. to count them, and we would know that no book has either left or come in since noon), boring as this truth may be. However, it is also known, in the circumstances, that, in actuality, no one has ever counted, is counting or will ever count how many books there are in my office at noon 28/10/2006. Therefore, it is known that it is not known that the number of books in my office at noon 28/10/2006 is 92, and so it is not epistemically possible that it is known that the number of books in my office at noon 28/10/2006 is 92. The previous

[31] Of course, for any candidate instance for extensional divergence, we cannot, on purely logical grounds, know that it is a ultimately suitable one (this being just another case of the structural unknowability so nicely illustrated by the Church-Fitch paradox): if it is known that it is not feasibly knowable that P, then it is known that it is not known that P (by the entailment from knowledge to feasible knowability and closure of knowledge), and so it is not known that it is not known that it is not known that P (by factivity of knowledge)—thus, it is not known that it is epistemically possible to know that P (by closure of knowledge). Hence the suppositional character of the counterexample against the sufficiency of epistemic possibility of knowledge for feasible knowability.

considerations concerning extensional and intensional non-equivalence still apply to this second kind of counterexample.[32]

Epistemic possibility of knowledge is thus neither necessary nor sufficient for feasible knowability. Therefore, the peculiar failures of transitivity we observed for the latter in the previous section cannot be explained by the non-transitivity (under failure of (KK)) of the former. Moreover, these considerations are already sufficient to block the third argument Brogaard and Salerno give against any restricted version of (EC). The argument assumes that the feasible-possibility operator behaves like an epistemic-possibility operator at least in the sense of satisfying:

(EP) If it is known that it is not the case that P, then it is not feasibly possible that P,[33]

and assumes $\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi$ and $\neg\varphi$ to be subject to (EC) if φ is. It runs as follows (adapted from [2, pp. 147–148]):

E_1	(E_1)	$\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi$	A
E_2	(E_2)	$\mathcal{K}(\varphi \equiv \blacklozenge\mathcal{K}\varphi)$	$\mathcal{K}((\text{EC}) \wedge (\text{D}))$ $\mathbf{IKT}_{\mathcal{K}}$
E_3	(E_3)	$\varphi \equiv \blacklozenge\mathcal{K}\varphi$	A
E_2	(E_4)	$\varphi \equiv \blacklozenge\mathcal{K}\varphi$	E_2 $\mathbf{IKT}_{\mathcal{K}}$
E_5	(E_5)	$\mathcal{K}(\mathcal{K}\neg\varphi \supset \neg\blacklozenge\varphi)$	$\mathcal{K}(\text{EP})$
E_6	(E_6)	$\mathcal{K}\neg\varphi \supset \neg\blacklozenge\varphi$	A
E_5	(E_7)	$\mathcal{K}\neg\varphi \supset \neg\blacklozenge\varphi$	E_5 $\mathbf{IKT}_{\mathcal{K}}$
E_1,E_2	(E_8)	$\blacklozenge\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	E_4 $[\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi/\varphi]$, E_1 $\mathbf{IP}$
E_9	(E_9)	$\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	A
E_9	(E_{10})	$\mathcal{K}\neg\mathcal{K}\varphi$	E_9 $\mathbf{IKT}_{\mathcal{K}}$
E_6,E_9	(E_{11})	$\neg\blacklozenge\mathcal{K}\varphi$	E_6 $[\mathcal{K}\varphi/\varphi]$, E_{10} $\mathbf{IP}$
E_3,E_6,E_9	(E_{12})	$\neg\varphi$	E_3,E_{11} $\mathbf{IP}$
E_9	(E_{13})	$\mathcal{K}\neg\mathcal{K}\neg\varphi$	E_9 $\mathbf{IKT}_{\mathcal{K}}$
E_6,E_9	(E_{14})	$\neg\blacklozenge\mathcal{K}\neg\varphi$	E_6 $[\mathcal{K}\neg\varphi/\varphi]$,$E_{13}$ $\mathbf{IP}$
E_3,E_6,E_9	(E_{15})	$\neg\neg\varphi$	E_3 $[\neg\varphi/\varphi]$,E_{14} $\mathbf{IP}$

[32] Of course, again, for any candidate instance for extensional divergence, we cannot, on purely logical grounds, know that it is a ultimately suitable one (this being yet just another case of the structural unknowability so nicely illustrated by the Church-Fitch paradox): if it is known that it is feasibly knowable that P, then it is known that P (by factivity of feasible knowability and closure of knowledge), and so it is not known that it is not known that P (by factivity of knowledge), and so it is not known that it is not the case that it is not known that it is not known that P (by factivity of knowledge)—thus, it is not known that it is not epistemically possible to know that P (by closure of knowledge and assuming that the consistency condition is sufficient for epistemic possibility). Hence, again, the suppositional character of the counterexample against the necessity of epistemic possibility of knowledge for feasible knowability.

[33] Notice that (EP) is intuitionistically equivalent with the necessity direction of the consistency condition (substituting ‘feasibly’ for ‘epistemically’ in ‘If it is epistemically possible that P, then it is not known that it is not the case that P’).

E_3,E_6,E_9	(E_{16})	$\neg\varphi \wedge \neg\neg\varphi$	E_{12},E_{15} **IP**
E_3,E_6	(E_{17})	$\neg\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	E_9,E_{16} **IP**
E_2,E_5	(E_{18})	$\mathcal{K}\neg\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	E_2,E_3,E_5,E_6,E_{17} $\mathbf{IKT}_{\mathcal{K}}$
E_2,E_5	(E_{19})	$\neg\blacklozenge\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	E_7 $[\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)/\varphi]$,$E_{18}$ **IP**
E_1,E_2,E_5	(E_{20})	$\blacklozenge\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi) \wedge \neg\blacklozenge\mathcal{K}(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	E_8,E_{19} **IP**
E_2,E_5	(E_{21})	$\neg(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	E_1,E_{20} **IP**

where φ is an arbitrary sentence falling under the scope of the restricted version of (EC). What (E_1)–(E_{21}) would show is that (even) a restricted (known) version of (EC) is committed to the rejection that neither φ nor $\neg\varphi$ are known. Although this needn't be a straightforward *reductio* of a restricted (known) version of (EC), it is worth noting that the conclusion (E_{21}) intuitionistically entails $\neg\mathcal{K}\varphi \supset \neg\varphi$ and classically entails $\varphi \supset \mathcal{K}\varphi$ (which are the main results of the original paradox of knowability), as the following simple proof demonstrates:

1	(1)	$\neg(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	A
2	(2)	$\neg\mathcal{K}\varphi$	A
3	(3)	φ	A
4	(4)	$\mathcal{K}\neg\varphi$	A
4	(5)	$\neg\varphi$	4 $\mathbf{IKT}_{\mathcal{K}}$
3,4	(6)	$\varphi \wedge \neg\varphi$	3,5 **IP**
3	(7)	$\neg\mathcal{K}\neg\varphi$	4,6 **IP**
2,3	(8)	$\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi$	2,7 **IP**
1,2,3	(9)	$\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi \wedge \neg(\neg\mathcal{K}\varphi \wedge \neg\mathcal{K}\neg\varphi)$	8,1 **IP**
1,2	(10)	$\neg\varphi$	3,9 **IP**
1	(11)	$\neg\mathcal{K}\varphi \supset \neg\varphi$	2,10 **IP**

(contraposition on (11), $\varphi \supset \neg\neg\varphi$ and the distinctively classical $\neg\neg\mathcal{K}\varphi \supset \mathcal{K}\varphi$ would then yield $\varphi \supset \mathcal{K}\varphi$). (E_1)–(E_{21}) would entail that a restricted (known) version of (EC) is no better off than the unrestricted one, and therefore has no point. However, (E_1)–(E_{21}) crucially relies on (known) (EP), and this has been shown not to hold generally by the second counterexample just offered.

18.6 Conclusion

We have seen how the factivity constraint on feasible knowability is crucial in generating peculiar failures of transitivity for the feasible-accessibility relation, and how such a failure cannot be explained, tempting as that may be, by an analysis of feasible knowability in terms of epistemic possibility of knowledge. I conclude that, quite unsurprisingly, the arguments we have been reviewing against a restricted version of (EC), when carefully scrutinized, prove unsound. They do so because they obliterate the distinctive semantics and logic—induced by the factivity constraint—which govern the modal epistemic operator ‘it is feasibly knowable that’ and simply take for granted that it is analyzable in terms of knowledge and some usual kind of metaphysical (or epistemic) possibility. Although this paper hasn’t been concerned with the fine details of such semantics and logic, it has indicated some important overall features that these can be expected to have, and has traced their source back to the the strain generated by the opposite requirements of *convergence* with actuality (as far as *what is known* is concerned) and of *divergence* with it (as far as the relevant *state of knowing* is concerned).

Acknowledgements Earlier versions of the material in this paper were given in 2005 at the Arché Modality Seminar (University of St Andrews); in 2006, at the International Colloquium on (*Anti-*)*Realisms, Logic and Metaphysics* (University of Nancy 2); in 2007, at the Arché Epistemology Seminar (University of St Andrews). I would like to thank all these audiences for very stimulating comments and discussions. Special thanks go to Ross Cameron, Roy Cook, Bob Hale, Jonathan Lowe, Agustín Rayo, Sven Rosenkranz, Crispin Wright and an anonymous referee. In writing this paper, I have benefitted, at different stages, from an AHRC Doctoral Award and from a RIP Jacobsen Fellowship.

References

1. Božić, M., and K. Došen. 1984. “Models for Normal Intuitionistic Modal Logics.” *Studia Logica* 43:217–45.
2. Brogaard, B., and J. Salerno. 2002. “Clues to the Paradoxes of Knowability: Reply to Dummett and Tennant.” *Analysis* 62:143–50.
3. Brogaard, B., and J. Salerno. 2009. “Fitch’s Paradox of Knowability.” *Stanford Encyclopedia of Philosophy*. Accessed March 4, 2010. http://plato.stanford.edu/entries/fitch-paradox/.
4. Došen, K. 1985. “Models for Stronger Normal Intuitionistic Modal Logics.” *Studia Logica* 44:39–70.
5. Dummett, M. 1975. *Truth and Other Enigmas*. London: Duckworth.
6. Dummett, M. 2001. “Victor’s Error.” *Analysis* 61:1–2.
7. Fitch, F. 1963. “A Logical Analysis of Some Value Concepts.” *The Journal of Symbolic Logic* 28:135–42.
8. Hacking, I. 1967. “Possibility.” *The Philosophical Review* 76:143–68.
9. Mackie, J. 1980. “Truth and Knowability.” *Analysis* 40:90–92.
10. Rosenkranz, S. 2004. “Fitch Back in Action Again?” *Analysis* 64:67–71.
11. Salerno, J. 2009. “Knowability Noir: 1945–1963.” In *New Essays on the Knowability Paradox*, edited by J. Salerno, 29–48. Oxford: Oxford University Press.
12. Tennant, N. 1997. *The Taming of the True*. Oxford: Clarendon Press.
13. Tennant, N. 2001. “Is Every Truth Knowable? Reply to Williamson.” *Ratio* 14:263–80.
14. Tennant, N. 2002. “Victor Vanquished.” *Analysis* 62:135–42.
15. Tennant, N. 2009. “Revamping the Restriction Strategy.” In *New Essays on the Knowability Paradox*, edited by J. Salerno, 223–38. Oxford: Oxford University Press.
16. Tennant, N. 2010. “Williamson’s Woes.” *Synthese* 173:9–23.
17. Williamson, T. 1992. “On Intuitionistic Modal Epistemic Logic.” *Journal of Philosophical Logic* 21:63–89.
18. Williamson, T. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.
19. Williamson, T. 2000. “Tennant on Knowable Truth.” *Ratio* 13:99–114.
20. Williamson, T. 2009. “Tennant’s Troubles.” In *New Essays on the Knowability Paradox*, edited by J. Salerno, 183–204. Oxford: Oxford University Press.

21. Zardini, E. 2008. “Truth and What Is Said.” *Philosophical Perspectives* 22:545–74.
22. Zardini, E. 2011. “Truth without Contra(di)ction.” Forthcoming in *The Philosophical Review*.
23. Zardini, E. 2011. “Truth, Demonstration and Knowledge. A Solution to the Paradox of Knowability.” ms.
24. Zardini, E. 2011. “Black Boxes. The Semantics and Logic of Obliterative Modalities”. ms.

Index

S. Rahman et al. (eds.), *The Realism-Antirealism Debate in the Age of Alternative Logics*, Logic, Epistemology, and the Unity of Science 23,
DOI 10.1007/978-94-007-1923-1, © Springer Science+Business Media B.V. 2012