Floris J. Bex

Law and Philosophy Library 92

# Arguments, Stories and Criminal Evidence

## *A Formal Hybrid Theory*

Springer

ARGUMENTS, STORIES AND CRIMINAL EVIDENCE

Law and Philosophy Library

VOLUME 92



For further volumes:
http://www.springer.com/series/6210

# ARGUMENTS, STORIES AND CRIMINAL EVIDENCE

## A Formal Hybrid Theory

by

Floris J. Bex

*University of Dundee, UK*

Springer

Floris J. Bex
University of Dundee
School of Computing
DD1 4HN Dundee
United Kingdom
florisbex@gmail.com

ISSN 1572-4395
ISBN 978-94-007-0139-7 e-ISBN 978-94-007-0140-3
DOI 10.1007/978-94-007-0140-3
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010938619



Printed on acid-free paper

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

# Preface

This book is based on my 2009 Doctoral dissertation “Evidence for a Good Story – A Hybrid Theory of Arguments, Stories and Criminal Evidence”. This dissertation was written while working on the project *Making Sense of Evidence*, in which a theoretically sound sense-making and visualization tool for Dutch police analysts has been developed. During the project I worked at the Centre for Law and ICT at the University of Groningen. I am grateful to my supervisors Henry Prakken, Bart Verheij and Peter van Koppen. I owe much to the knowledge and advice they have shared with me over the years. I also thank John-Jules Meyer, Arend Soeteman and William Twining for taking the time to read, assess and comment on my dissertation. Furthermore, I thank Susan van den Braak, Gerard Vreeswijk and Herre van Oostendorp, who worked on the project at the University of Utrecht and who have provided important feedback by implementing and testing the sense-making system AVERS.

I thank everyone who, over the years, has shown an interest in and commented on my work. The argumentation, evidence and the AI and Law communities I thank for the interesting and fruitful conferences and discussions. In particular, I would like to mention Katie Atkinson, Trevor Bench-Capon, Tom van Engers, Tom Gordon, Jeroen Keppens, Chris Reed, Burkhard Schafer and Douglas Walton.

Dundee, UK Floris J. Bex
July 2010

# Contents

# Chapter 1
# Introduction

The subject of this book is reasoning with evidence to establish the facts in criminal cases. In a legal context, the study of evidence is often equated with the study of the *law of evidence*, for example, the legal rules of evidence that govern which types of evidence are legally valid or admissible.[1] However, a large part of the study of evidence, and particularly reasoning with evidence, constitutes the study of the *rational process of proof*. This process involves reasoning with observed evidence and commonsense knowledge of the world around us in order to establish whether something is or was the case, that is, to establish the facts of the case. The rational process of proof and the reasoning employed in this process is the central theme of this book.

## 1.1 Rational Theories of the Process of Proof

The distinction between the study of evidence law and the study of the process of proof was made in the beginning of the twentieth century by Wigmore, as follows: "The study of the principles of evidence [. . .] falls into two distinct parts. One is Proof in the general sense – the part concerned with the ratiocinative process of contentious persuasion – [. . .]. The other part is Admissibility– the procedural rules devised by the law" (Wigmore, 1931, p. 3). In the process of proof, the reasoning takes the form not of legal reasoning, but rather of commonsense reasoning: viz. "the counsel sets himself the task [. . .] of persuading the jury that they should or should not believe the fact [. . .]. To do this, he must reason naturally, as all men reason [. . .]" (Wigmore, 1931, § 1–2). Wigmore argued for the development of a "science of judicial proof". This science of proof should formulate rational principles for reasoning with evidence and proof independent of the rules of law. He himself set out to develop such a rational theory which could be used for structuring and analysing arguments based on a mass of evidence.

[1] Examples of rules about the legal validity or admissibility of evidence can be found in the Dutch Code of Criminal Procedure (DCCP) and the American Federal Rules of Evidence (FRE), respectively.

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_1, 

Twining also distinguishes between legal and rational, commonsense features of the study of evidence in his discussion of the Rationalist Tradition of Evidence Scholarship.[2] The fundamental principle of the Rationalist Tradition, which according to Twining has served as the basis for most of the Anglo-American research on judicial evidence of the last 250 years, is that the main objective of procedural law is to correctly apply the law to facts which are considered to be proven and that these facts should be proven by rational means. The assumptions that underlie the Tradition can be expressed in two models: a rationalist model of adjudication and a model containing assumptions for a rationalist theory of proof. This second model assumes that knowledge about particular past states and events in the world is possible and that it is necessary to rationally reason with the evidence in order to establish whether or not our belief in the truth of such events is justified.

It is this rational and commonsense "pursuit of truth" that is the main interest of this book and the assumptions contained in the rationalist model of proof are also at the basis of the current work. More specifically, the current view on rationality can be characterized as *bounded procedural rationality* (Rescher, 1977; Simon, 1982): a belief or decision is rational if it is in agreement with the knowledge that has actually been considered (or should have been considered) in a proper procedure. This rationality is bounded because humans are limited by cognitive and practical constrains in their consideration of knowledge; it is procedural because the rationality of a belief depends on the quality of the process that has been used to obtain the belief. Note that the rational process of proof is by no means exclusive to a legal setting. In science but also in our everyday lives we continually use and reason with evidence in order to infer conclusions; Schum speaks of "the context in which evidence arises", for example scientific, medical or legal evidence (Schum, 1994, p. 1).

After Wigmore, the interest in a rational theory of proof decreased. Most legal theorists continued their research which focused more on the model of adjudication rather than on factual inference. However, in the past 30 years, a movement that has become known as the "New Evidence Scholarship"[3] has become an important force in research on evidence. New Evidence Scholarship, which is firmly based on the Rationalist Tradition, includes researchers who have the same interests, namely factual evidence and reasoning with this evidence in a legal context. Subjects are, among others, the logic of inferences about facts and how to use formal probabilistic methods in evaluating evidence. New Evidence Theorists such as Anderson, Schum, Tillersand Twining explicitly build on Wigmore's ideas in their development of *Modified Wigmorean Analysis*.[4] Central in this analysis is the reasoning from the

[2]Most of Twining's work on evidence includes a section on the Rationalist Tradition. For an overview see (Anderson et al., 2005, pp. 78–86) and for a more extensive account see Twining (1994, Chapter 3) and (2006, Chapter 3).

[3]The term was coined by Richard Lempert (1986)

[4]This term was to my knowledge first used by Twining (2007). The research on Modified Wigmorean Analysis encompasses a large amount of interesting material on a multitude of different subjects. The "locus classicus" would be Wigmore's (1931) work. Important new work in

evidence to the propositions that have to be proven and the use of detailed graphs to logically structure and analyse this reasoning from evidence to conclusions. The authors argue that this type of analysis can be useful in both the investigative and the decision-making phase, as subjecting the evidence and reasoning in a case to a thorough Wigmorean analysis allows one to identify sources of doubt. Thus possible miscarriages of justice – an example is the famous Sacco and Vanzetti case – as well as mistakes in criminal investigation – Anderson and colleagues mention that the mistakes made by the intelligence services surrounding the 9/11 events – can perhaps be prevented.

In contrast to the New Evidence Theorists in the Anglo-American law community, the Dutch legal community does not have a long and broad tradition of research into the rational and non-legal aspects of proof.[5] However, in the past decades, interest in the psychological background of reasoning with evidence has grown. This growing interest is partly prompted by a number of (perceived or possible) miscarriages of justice which have not been caused by wrongly interpreting or applying the law, but rather by mistakes of a psychological nature. For example, in some cases the police investigation suffered because of the well-known effect of confirmation bias: a tendency to search for and interpret evidence in a way that conforms to one's prior beliefs (e.g. a suspect's guilt), while dismissing evidence that might point to other hypotheses (e.g. the suspect's innocence).

In 1993, the legal psychologists Crombag, van Koppen and Wagenaar published an influential book entitled *Dubieuze Zaken – de psychologie van strafrechtelijk bewijs* (Dubious Cases – the psychology of criminal evidence). In this book, they discuss by means of example cases a number of possible mistakes people make when reasoning with evidence and proof.[6] They also provide a rational and normative theory of reasoning with criminal evidence. This theory takes as its basis earlier empirical research and theories by Bennett and Feldman (1981) and Pennington and Hastie (1986, 1992, 1993b), where it is claimed that investigators and jurors use stories about "what happened" in a case to organize and analyse the available evidence. In order to sidestep the problem of believing a "good story" above a "true story", Crombag, van Koppen and Wagenaar's Anchored Narratives Theory postulates that stories should be anchored in commonsense knowledge of the world around us.

this school of thought is the *Analysis of Evidence*, authored by Anderson, Twining and Schum (2005). This book, which I have used extensively in the development of my own ideas, contains information on a wide variety of subjects regarding reasoning with evidence and proof. Other seminal work in the tradition of the New Evidence Theory is contained in *Rethinking Evidence* (Twining, 2006). Finally, Peter Tiller's website (Tillers, 2006) can also be used as an invaluable source on (the law of) evidence in the tradition of the New Evidence Theory.

[5] Nijboer (2000, p. 28), however, argues that investigation and proof in criminal cases can be characterized as "special forms of empirical investigation and proof" and that Dutch jurists largely agree with a rationalist notion of knowledge.

[6] An English adaptation was publishedas (Wagenaaretal.,1993). Other work which stands in the same tradition is Wagenaar and Crombag (2005) and De Poot et al. (2004), in which the Anchored Narratives Theory is applied to police investigation.

## 1.2 Making Sense of Evidence

The above-mentioned theories on reasoning with evidence and proof almost all have a descriptive as well as a normative side: on the one hand, the theories try to model the patterns of reasoning that are employed in the process of proof and on the other hand, the theories try to indicate the shortcomings of people and mistakes that are made when reasoning with masses of evidence and propose ways in which these mistakes can be avoided. In a way, theories for reasoning with evidence specify how we can and should *make sense of evidence*. In other words, given a mass of evidence how can we best structure and represent the evidence?

In any (larger) case it is highly important that the lines of reasoning and the evidence, hypotheses and background knowledge used in the reasoning are made explicit. In this way, sources of doubt in the reasoning can be identified and reasoned about. Furthermore, explicitly identifying and structuring all hypotheses lessens the danger of so-called *tunnel vision*, where the most likely scenario is taken as the leading hypothesis and alternatives are insufficiently considered.

Various tools, such as tables, stories or visualisation aids can be used in the sense-making process. For example, Anderson and colleagues (2005) argue that charting the reasoning from evidence to conclusions is necessary in order to expose sources of doubt in the reasoning. Wagenaar and colleagues (1993) and Pardo and Allen (2007) argue that stories are a natural tool humans should use when talking about a mass of evidence and Heuer's (1999) procedure for analysing hypotheses contains a step in which the various alternatives are ordered in a matrix. A relatively new development concerning sense-making and (criminal) evidence is the emergence of computer-based support tools for investigators and decision makers. Such a support tool is a computer program that allows for the electronic management of evidence and scenarios in a case.[7] Through a combination of spreadsheets and (timeline) visualization functions, these tools allow the user to give an overview of the evidence and scenarios in a case and link the evidence to specific persons or places.

Making sense of evidence using the various tools is important in all stages of the process of proof. In the investigative phase, the amount of evidence and hypotheses can quickly grow and various representations of the hypotheses and evidence serve as reminders and facilitate the communication between the investigators. Thus, in the investigative phase the sense-making tools are used for the general purpose of keeping track of all the incoming information and the reasoning associated with this information. In the decision-making phase, sense-making is oriented towards a specific goal. For example, Anderson and colleagues argue lawyers can use visualisation techniques to identify weaknesses in their own arguments and in arguments from the opposing party. In an adversarial system, this allows a party to strengthen his own case by anticipating counterarguments and weaken the opponent's case by

[7] Examples are CaseMap (http://www.casesoft.com/casemap/casemap.asp; accessed on 26 July 2010) and Analyst's Notebook (http://www.i2.co.uk/Products/Analysts_Notebook/default.asp; accessed on 26 July 2010).

attacking them at their weakest point. Nijboer and Sennef (1999) argue that the justification for a decision that judges give in the Dutch inquisitorial system should not only be aimed at allowing higher authorities to check and control the decision on its legal merits but also make the decision about the facts of the case and its justification understandable to the general public.[8]

## 1.3 Reasoning with Evidence in Artificial Intelligence and Law

Whilst the general view on rational reasoning with evidence as described above stems mainly from legal theory and legal psychology, the basis of this book is firmly in the field of *Artificial Intelligence and Law*. Artificial Intelligence (AI) is a multidisciplinary field which combines insights from diverse disciplines such as cognitive psychology, computer science and philosophy. Because of this multidisciplinary background, logical models of knowledge and reasoning in AI are constructed with differing aims in mind. For example, a model of reasoning may be intended as a cognitive model that describes actual human reasoning, as a computational model that forms the basis of programs for automated reasoning or as a theoretical model that conceptually analyzes knowledge and reasoning.[9] For a computational model the advantage of a logical model is that such a model can be more readily understood by a computer than models expressed in, for example, natural language.[10] In descriptive cognitive or theoretical models, a formal logic adds a level of detail and specificity that can take away ambiguities and thus allow for the detection of omissions, errors and inconsistencies.

In the field of AI and Law, insights from general AI are applied to topics which are typically studied in law and legal theory; the reasoning that is formally modelled in AI and Law hence concerns legal reasoning. For example, Loui and Norman (1995), Prakken and Sartor (1996) and Verheij, Hage and van den Herik (1998) model legal rules in a formal logic and Ashley (1991) provides formal models of legal cases. Most of the formal research in AI and Law focuses on reasoning with legal rules and cases. However, in the past decade the interest in formal theories of reasoning with evidence and crime scenarios has also emerged. For example, Verheij (2000) compares the Anchored Narratives Theory to formal logics for argumentation and Bex, Prakken, Reed and Walton (2003) have modelled Wigmore charts using a formal argumentation logic. Keppens and colleagues (e.g. Keppens and Schafer, 2006) provide a logical model-based approach to reasoning with crime scenarios. There are also formal approaches to modelling reasoning with evidence which do not use a symbolic logic but rather a quantitative approach. Thagard (2004,

[8]The increased public nature of decisions in the Netherlands (cf. van Lent, 2008) forces judges to explain their decisions more thourougly and intelligibly.

[9]These three aims are adapted from Verheij's (1996) discussion of the aims of formally modelling argumentation.

[10]See Prakken (1997, Chapter 1) for a brief discussion of the role of logic in AI.

2005) models stories and evidence in connectionist "coherence networks" and statistically inspired Bayesian Network theories of reasoning with Wigmore charts have, for example, been studied by Kadane and Schum (1996), Levitt and Laskey (2000) and Hepler and colleagues (2007).

Within AI and Law, formal models of legal reasoning or reasoning with evidence might serve any of the above-mentioned aims of logical models in AI. For example, Prakken (1997) notes that a concern of AI and Law research is to provide formal rational theories that act as foundations for computer programs. An example is Thagard's theory of explanatory coherence, which has been implemented as a system for automated reasoning called ECHO (Thagard, 1989). Hage (1996) argues for a theoretical logical model of legal reasoning when he says that logic can be an intermediary between on the one had a jurisprudential account of legal reasoning and on the other hand a computational account of legal reasoning. Such a logic abstracts from the characteristics of a specific legal system and also does not aim to give a model that is necessarily computationally feasible. Verheij's (2000) interpretation of the Anchored Narratives Theory and Bex and colleagues' (2003) treatment of Wigmore charts fall in this category.

In addition to computer systems that reason automatically, such as classic knowledge-based expert systems, the interest in *sense-making systems* has recently also grown in AI (and Law). Sense-making systems do not contain a knowledge base and do not reason automatically, but instead help the user make sense out of a certain problem by allowing the user to logically structure his knowledge. One development in this respect is the emergence of argument visualization tools.[11] Based on ideas from critical thinking and argumentation theory these tools allow the user to structure and visualize the reasoning employed in a case according to some specific underlying logical theory of reasoning.[12] This underlying theory serves multiple aims. One aim is to enforce a standard of rationality by requiring that the user's reasoning stays within the logical system. Another use of an underlying theory is that the sense-making tool can perform some computations; for example, showing the user which arguments can be accepted and which should be rejected according to the current assumptions.

A logical theory underlying a sense-making system should essentially find a middle ground between the three aims of logical models (i.e. conceptual, cognitive and computational). In order to provide a standard of rationality, a theoretical model should precisely define the various core concepts that apply to the particular mode of reasoning. Because of the aim of sense-making, however, the model should also fit with cognitive models of reasoning so as to ensure that it employs concepts that

---

[11] Examples are Araucaria (http://araucaria.computing.dundee.ac.uk; accessed on 26 July 2010) and Rationale (http://rationale.austhink.com; accessed on 26 July 2010). See (Verheij, 2005b; van den Braak, 2010, pp. 35–45) for overviews.

[12] In some tools, like Rationale, the underlying logic is basic and largely implicit (cf. van Gelder, 2007) whereas other tools, like Argumed, essentially allow the user to build arguments using an explicit argumentation logic (cf. Verheij, 1999).

are natural to an everyday user, which cannot be expected to have in-depth knowledge of formal models of reasoning. Finally, the model should have a computational side so that the tool can help the user by performing computations.

## 1.4 Research Goals

Section 1.1 discusses the idea of rational theories of criminal evidence. In the research on such theories, essentially two trends can be distinguished. The research by the New Evidence Theorists such as Anderson, Schum, Tillers and Twining largely focuses on the use of detailed Wigmorean argument charts to structure and analyse a mass of evidence and to expose sources of doubt in the reasoning.[13] In contrast, Crombag, van Koppen and Wagenaar's Anchored Narratives Theory uses stories to organize and analyse available evidence. The two methods share many ideas and the central concepts of argument and story play an (often implicit) role in both methods. The New Evidence Theorists's Modified Wigmorean Analysis is complemented by outlines, chronologies and stories. For example, the *Analysis of Evidence* (Chapters 6 and 10) and *Rethinking Evidence* (Chapters 10 and 11) discuss several aspects of stories and the use of stories for providing an overview of a case, for identifying gaps in a case and for making a persuasive case in court. Although Crombag, Van Koppen and Wagenaar focus on the story-based perspective in both in their choice of wording and in their research background, several of their central claims have a more argumentative than story-based flavour. Especially the role of generalizations (or anchors), exceptions to these generalizations and of the dynamics of developing and refining an analysis of the evidence in a case are characteristic for the argumentative slant of the approach by Crombag, Van Koppen and Wagenaar.

Despite the appearance of arguments and stories in Modified Wigmorean Analysis and the Anchored Narratives Theory, none of these theories fully integrates stories and arguments in one concise rational theory of reasoning with evidence. For example, stories and their rational analysis can and should play a bigger part in the analysis of a mass of evidence; in particular, their precise role in generating hypotheses and finding "gaps" is at present not clarified in Modified Wigmorean Analysis. In the Anchored Narratives Theory, the evidential data has no clear place and the various ways of argumentative reasoning from evidential data to a conclusion are not discussed in detail.

The main goal of this book is to propose a hybrid argumentative story-based theory which combines reasoning with arguments and stories. An informal as well as a formal logical version of the theory will be developed, in an attempt to make the core ideas and concepts accessible to a wide audience consisting of lawyers, legal theorists, psychologists and formally oriented researchers from AI and Law. A general conceptual framework will be constructed, in which both reasoning with

---

[13] A notable exception here are Pardo and Allen (2007), who advocate using stories to explain the evidence.

stories and arguments will be discussed. In this discussion, insights from various fields of research[14] will be combined with new ideas to form an informal hybrid theory of reasoning with evidence. Aside from the academic pursuit of developing a theoretical account of reasoning with evidence, the hybrid theory is also intended to provide lessons for the investigative and judicial practice by formulating guidelines or heuristics for correct and rational reasoning in various contexts of investigation and decision making. Modelling the hybrid theory in a formal logic forces the precise and detailed definition of the various ideas proposed in the informal theory and compels us to make clear choices as to how the various concepts should be defined.

The need for the development of a logical theory is emphasized since it facilitates the implementation of the theory in a sense-making tool. In *Making Sense of Evidence*,[15] the coordinating research project of this work, a theoretically sound sense-making and visualization tool for Dutch police analysts has been developed. This tool, called AVERS (Argument Visualization for Evidential Reasoning based on Stories, see van den Braak, 2010; van den Braak et al., 2007; Bex et al., 2007b), combines reasoning with arguments and scenarios and is based on the logical theory as developed in this book. As was discussed above, a logical model or theory that serves as the basis for a sense-making system should meet essentially three criteria. First, the theory should be *natural* in that it employs concepts that are natural to an everyday reasoner such as a crime analyst or a judge. Investigators and decision makers cannot be expected to have in-depth knowledge of mathematical or formal models and therefore the theory should be based on reasoning forms used in practice. Second, the theory should be *rationally well-founded*, that is, the theory should provide a clear rational framework which is in agreement with the prevailing (legal and philosophical) theories of rational reasoning about evidence. Furthermore, the theory should ideally encourage a correct and general standard of rational reasoning. Like in all kinds of commonsense reasoning, mistakes can be made in reasoning with criminal evidence; because such mistakes can have a large impact on the life of a person (e.g. conviction of an innocent) or society as a whole (e.g. acquittal of a murderer), they should be avoided. Given the current conception of procedural rationality, the theory should facilitate and promote a proper procedure for performing an inquiry concerning evidence in a criminal case, thus promoting rational reasoning about the evidence. Finally, the theory should be *formally specified* with an eye towards software development, so that it can act as a proper foundation for AVERS. This means that the model of reasoning ideally has a computational side so that it can aid the investigators by, for example, computing which possible scenario is best supported by evidence. However, since only a small amount of automated

---

[14] In particular legal theory, legal psychology, philosophy, argumentation theory, cognitive modelling and AI.

[15] The project is a collaboration between the Centre for Law and ICT (University of Groningen), the department of Artificial Intelligence (University of Groningen) and the Intelligent Systems Group (Utrecht University). For more information, see: http://www.cs.uu.nl/research/projects/evidence/ (last accessed on July 19, 2010).

reasoning is performed by AVERS, the hybrid theory does not necessarily have to be computationally feasible.

To summarize, the research questions of this book can be formulated as follows:

- What form does a theory of reasoning with criminal evidence have?
  - What are the roles of arguments and stories in reasoning with evidence and how can they be combined into a hybrid theory?
  - What are the necessary elements in a theory that is natural and rationally well-founded?
  - Given the current notion of procedural rationality (p. 2), how can some standard of rationality for reasoning with evidence be set?
- How can reasoning with stories and arguments be specified in a formal logic? Conversely, what do insights from formal logic teach us about reasoning with stories and arguments?
- What form does a formal procedure and protocol for building a logical theory of stories and arguments have?

The method through which these questions will be answered involves first the development of a broad conceptual theory based on various insights from the literature. This conceptual theory anticipates a formalization by already employing concepts and ideas from formal theories of reasoning. Furthermore, the conceptual theory is rendered visually in a "semi-formal" style that closely matches that of AVERS. The theory will then be formalized and any new ideas or errors or inconsistencies that emerge through the process of formalization will be incorporated or changed in the general theory. Finally, a case-study is performed in which an actual case is modelled in the hybrid theory. This case study is intended as a test of the hybrid model in that it shows the possibilities and limitations of the theory when analysing a mass of evidence in a large case.

## 1.5 Book Outline

The structure of this book largely follows the method of research as described above. Chapter 2 is a general introduction to criminal evidence and the process of proof. In this chapter, the various basic concepts, processes and modes of reasoning that apply to the current context are introduced. In Chapter 3 arguments and stories and their respective roles in reasoning with evidence are discussed. This discussion will take the form of a comparison of two separate approaches to reasoning with evidence, the argument-based and the story-based approach. These approaches are each not based on one particular author or group of authors, but abstract accounts of the two trends in research about evidence as discussed above. Consequently, Sections 3.1 and 3.2 should not be regarded as a comparison of Modified Wigmorean Analysis and the Anchored Narratives Theory, but rather as a broad discussion the aim of which is to clarify the roles of arguments and stories in the process of proof. In

Chapter 4 the informal hybrid theory is developed and the way in which arguments and stories from the previous chapter can be combined is explored. Chapter 5 contains the logical hybrid theory, in which the ideas from Chapters 3 and 4 are further formally elaborated. As with the informal theory, arguments and stories are first treated separately in Sections 5.2 and 5.3, respectively. Then the formal hybrid theory is proposed analogously to Chapter 4. In Section 5.5, the protocol for a rational inquiry dialogue about evidence is proposed. Chapter 6 contains the case study in which the theory is applied to an actual case. Chapter 7 describes relations with other research and Chapter 8 concludes the book.

# Chapter 2
# Reasoning with Criminal Evidence

In this book, the main subject is the study of the rational *process of proof*. This process involves reasoning with observed evidence and general commonsense knowledge of the world around us in order to establish the facts of the case. This chapter discusses the basic assumptions and concepts in the process of proof. Because the research on reasoning with (legal) evidence is conducted in separate fields (i.e. general jurisprudence, philosophy (of law), psychology (of law), artificial intelligence), confusion sometimes arises over the meaning and precise definition of concepts used by the authors in the various fields. For example, a logician thinks of the concept "rule" in a different way than a lawyer think of the concept: where the logician might define "rule" as, for example, an inference rule (e.g. modus ponens), a lawyer might think of a rule as being a rule of law (e.g. "theft is punishable by *n* years imprisonment"). Furthermore, some of the terms can have different meanings: an "argument" can be a reason or set of reasons given in support of a proposition or a verbal fight or debate between people with differing views. The goal of the rest of this chapter is to provide some clarity as to what is meant with the various terms used in reasoning with evidence and the process of proof. In Section 2.1, three basic concepts – facts, evidence and world knowledge – will be discussed. Section 2.2 outlines the process of proof in general terms by describing the various phases in the process and Section 2.3 discusses some general conceptual background on the types of reasoning and inference in the process of proof. The ideas in these sections can be viewed as general ideas on evidence and proof and are largely applicable to any context in which reasoning with evidence is performed. However, some legal specifics will be mentioned in the text and the examples also have a distinct legal flavour.

## 2.1 Facts, Evidence and General Knowledge

In the rational process of proof, the *facts of the case* have to be proven by *evidence*. That the facts follow from the evidence is often by no means self-evident; in a complex case, using the evidence to prove the facts requires us to reason with this evidence. Such reasoning can be regarded as commonsense reasoning as, in addition to the knowledge gained from the case-specific evidence, we also use *general*

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_2, 

*commonsense knowledge* of the world in our reasoning. The notions of fact, evidence and general knowledge are central to the process of proof and will therefore be briefly introduced in this section.

### *2.1.1 Facta Probanda and Facta Explananda*

The *facts of the case* often denote the events or states of affairs that are assumed, at least for the moment, to have happened or existed (Anderson et al., 2005, p. 382). Following the common definition of *fact* in philosophy, Hage and Verheij (1999) define a fact as an obtaining state of affairs in the world expressed by a true proposition. Examples of facts are that John F. Kennedy was the 35th president of the United States and Jack Ruby shooting Lee Harvey Oswald. States of affairs do not have to obtain; for example, the state of affairs that Liverpool is a city in the Netherlands clearly does not obtain and hence the proposition "Liverpool is a city in the Netherlands" is false and is not a fact. In evidential reasoning, it is often impossible to definitively establish whether a state of affairs actually was (or is, or will be) the case, so the truth of a proposition cannot always be established with absolute certainty. It is for this reason that the propositions expressing such states of affairs are qualified, for example "it *may* rain tomorrow" or "OJ *might have been* Nicole Simpson's killer".

The facts of the case can also denote propositions the truth of which is unknown. Such propositions are often equated with the facts that need to be proven, the *facta probanda* (Wigmore, 1931, p. 8; Anderson et al., 2005, p. 384), or simply the *probanda* ("that which has to be proven").[1] Anderson and colleagues (2005) distinguish between ultimate, penultimate and interim probanda. The *ultimate probandum* is the major factual proposition that is at issue in a case, that is, the proposition that the prosecution needs to prove in order to have the defendant convicted. So in a murder case it must be proved that the defendant intentionally killed the victim. The ultimate probandum can usually be broken down into separate propositions, the *penultimate probanda*, which signify the individual elements of the ultimate probandum.

Another type of facts in the process of proof are the *facta explananda* (or simply *explananda*, "that which has to be explained"). These explananda play a major role in the investigation stage of the criminal process, when the police is looking for evidence and pursuing various avenues of investigation. Faced with some states or events that are out of the ordinary and that point to the possibility of a crime having been committed, the police investigators will try to explain these initial clues or explananda. The distinction and interaction between explanation and proof (i.e. between explananda and probanda) is central to the process of proof and will be further discussed in Section 2.2. For now, the term *fact* will stand for obtaining states

[1] This interpretation of the term fact is evident in the Dutch Code of Criminal Proceedings, which states that "the court deliberates [. . .] whether it is proved that the facts were perpetrated by the suspect [. . .] " (article 350 DCCP).

of affairs expressed by a justified proposition and propositions which need to be explained or proven will be called *explananda* or *probanda*.

## *2.1.2 Evidence*

The term *evidence* stands for the available body of information indicating whether a belief in some proposition is justified. When people talk about "the evidence in a case" they usually mean the *evidential data* (Anderson et al., 2005, p. 382), that is, the primary sources of evidence. This evidential data, here also called *items of evidence*, *pieces of evidence* or *sources of evidence*, comes to the reasoner through a process which Wigmore calls "autoptic proference" (Wigmore, 1931, p. 10), which means so much as perceiving with one's own senses. Notwithstanding sensory defects on the part of the reasoner, the existence of the evidential data itself cannot be sensibly denied. In other words, if we hear a testimony of a witness who says that he saw someone who looks like the suspect jump into a red car, the proposition expressing the state of affairs that the evidential datum exists – "there is a testimony by a witness who saw someone who looks like the suspect jump into a red car" – can be accepted as justified.

In addition to evidential data, the term evidence can also point to other propositions which have been accepted as true and which might have been inferred from evidential data themselves. For example, the proposition that "someone who looked like the suspect jumped into a red car", which has been inferred from the above testimony, is itself evidence for the proposition that "the suspect jumped into a red car". It is important that the evidential data and the propositions that are inferred from it are not confused. Anderson et al. (2005, p. 60) denote this as follows: $E^*$ stands for evidential data about event $E$; in the above example, $E^*$ is the testimony itself and $E$ is the event that "someone who looked like the suspect jumped into a red car". As was already noted, this event $E$ can then be evidence for another event $F$: "the suspect jumped into a red car". The separation between an event and the evidential data from which the event is inferred is important because the existence of the evidential data does not mean that the event actually happened. In the above example, the witness may be lying or he may misremember or the person who looked like the suspect may not be the suspect at all. In sum, the term *evidence* stands for the information that (positively or negatively) influences our belief about a particular proposition. This information can be a piece of evidential data but also a proposition which itself has been inferred from data.

In common law, a distinction is often made between *direct* and *circumstantial evidence*. Direct evidence, which is sometimes also called real evidence, is usually regarded as "evidence which, if believed, proves the existence of the fact in issue without inference or presumption".[2] Often, direct evidence is equated with witness testimony: "evidence which immediately points to the question at issue,

[2] *State v. Famber*, 358 Mo. 288, 293, 214 S.W.2d 40, 43.

or is evidence of the precise fact at issue and on trial, by witnesses who can testify that they saw the act done, or heard the words spoken which constitute the facts to be proved".[3] Circumstantial evidence, on the other hand, is "evidence from which the judge or jury may infer the existence of a fact in issue but which does not prove the existence of the fact directly".[4] This distinction between direct and circumstantial evidence has met with criticism: according to Anderson and colleagues (2005, pp. 76, 77) there is no such thing as direct evidence, as reasoning from evidential data $E^*$ to an event $E$ always involves at least one inference. Instead, they distinguish between directly relevant evidence and indirectly relevant or ancillary evidence (Anderson et al., 2005, pp. 62, 63). Evidential data $E^*$ is *directly relevant* for some proposition $P$ if a chain of reasoning can be constructed directly from the evidence to the proposition. Hence, in the above example of the man getting into the car the data $E^*$ is directly relevant for both $E$ and $F$. *Ancillary evidence* with respect to a proposition $P$ is only indirectly relevant for determining whether $P$ is justified; Anderson and colleagues define it as "evidence about other evidence and its probative strength" (Anderson et al., 2005, p. 380). There are two types of ancillary evidence: strengthening and weakening. An example of strengthening ancillary evidence in the above example would be that the witness is known as an honest person. An example of weakening ancillary evidence is that the witness is partially blind so he could not have seen the man get into the car.

The evidential data itself can be considered as the most important basis of the whole process of proof; it therefore pays to briefly discuss the various types of data. There are many types of data, for example, testimonies by witnesses or experts, documents, photographs, videos, objects and so on. Anderson and colleagues (2005, pp. 63–70) argue that there are two main types of evidence: testimonial and tangible evidence. Testimonial evidence concerns testimonies by witnesses, experts, suspects etcetera and tangible evidence concerns all other things that can count as evidence for something (e.g., objects, documents, videos). In the next paragraphs, some of the more common types of evidence and their features will be discussed.

*Witness testimony* is one of the most important types of evidential data. Because of its importance in evidential reasoning, it has been studied extensively in psychology (Loftus, 1996; Crombag et al., 1994, Chapters 9, 10, 11 and 14; Wagenaar et al., 1993, Chapter 8) and has a prominent place in general theories on evidence (Wigmore, 1931, Part III; Anderson et al., 2005, pp. 65–70). There are various kinds of witnesses and for each there are a significant number of ways of examining, hearing and interrogating the witness. For example, children in a sexual abuse case have to play with anatomically correct dolls, eyewitnesses are sometimes asked to point out the suspect in a line-up and witnesses who are also a suspect are interrogated differently from witnesses who are a victim. With regard to a testimony, there are a

[3] *Stern v*. Employers' Liability Assur. Corp., Ltd. Of London, England, 249 S.W. 739, 741.

[4] "circumstantial evidence", *Oxford Dictionary of Law*. Ed. Elizabeth A. Martin and Jonathan Law. Oxford University Press, 2006. *Oxford Reference Online*. Oxford University Press. Accessed on 15th July 2008. http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t49e613

number of important issues to consider. Was the witness in a position to know what happened? Does the witness give a simple account of what he saw or does the witness hypothesize a course of events? Is what the witness says internally consistent and consistent with other evidence? Is the witness biased in some way; for example, does he have something to gain by testifying in a certain way? Following original work by Schum (1994, p. 325), Anderson et al. (2005, pp. 66–68) give a summary of the various attributes of the credibility of witness testimonies. They have identified three main features of witness testimonies: (1) veracity, whether the witness himself believes what he says, that is, whether he is lying (2) objectivity, whether the witness correctly remembers what he observed, and (3) observational sensitivity, whether the witness correctly observed what happened. Several ways to strengthen or weaken the credibility of a witness testimony are given. Some of these are linked to a specific attribute; for example, if a witness has sensory defects his observational sensitivity is weakened. Other ways of strengthening or weakening are less specific; for example, a contradictory testimony of another witness may weaken our belief in one witness but it does not weaken a specific attribute.

To some extent, the law dictates in which ways witness testimonies may be used. For example, the Dutch Code of Criminal Procedure expressly states that a witness testimony contains only "facts and circumstances, which [the witness] himself perceived or experienced" (article 342 DCCP) and the Federal Rules of Evidence state that a witness may only testify to a matter of which the witness has personal knowledge (Rule 602 FRE). *Hearsay* evidence, when a witness testifies to something he did not experience himself but rather learnt about from someone else, is generally not admitted in many jurisdictions (e.g. Rule 802 FRE). In the Netherlands hearsay evidence is admitted as per a ruling of the Court of Cassation from 20th of December 1926 (NJ 1927, p. 85, "de auditu"), particularly because this allows for the use of documents drafted by police officers which contain witness testimonies. Testimonial evidence that is not admissible or legally valid can still be interesting for the process of proof. For example, witnesses who clearly accuse or exonerate a suspect may very well be biased and this should be taken into account when assessing the total body of evidence. Anderson and colleagues (2005, p. 66) remark that hearsay evidence can always be used in the investigation phase, where a hearsay testimony might lead to the person who witnessed the crime firsthand.

Another form of testimonial evidence is *expert testimony* (see Anderson et al., 2005, p. 270; Wagenaar et al., 1993, Chapter 9; Hielkema, 1999). Much of the information that is used in fact investigation and when making a decision in court is based on specialist knowledge that only an expert knows. For example, an average policeman, judge or juror does not have detailed knowledge of the specific techniques involved in DNA-analysis and needs to be told if and why two DNA samples match by a forensic specialist. In some respects, expert testimony is different from normal witness testimony. For example, while it is often not desirable that an eyewitness gives his opinion on matters (he should simply recount what he saw), the very nature of an expert testimony makes it an opinion. We ask an expert to interpret certain data for us and such an interpretation involves the expert inferring conclusions from the data he is provided with. In fact, because of the complex issues an expert testifies to we expect him to voice his opinion in simple terms, lest there be the danger of misinterpretation of the expert's conclusion by the investigator or decision maker.

Investigators and decision makers should carefully consider expert testimonies and expert evidence because expert witness can make the same, intentional or unintentional, mistakes as any person. For example, the expert may very well be biased. Furthermore, the expert might interpret certain findings in the wrong way or the process which the expert used to gather and test the data may be of dubious quality. Because the subject of the expert testimony is often too complex for the investigator or decision maker to fully discuss himself, an important check is to determine what the expert says is consistent with what other experts in the field say. In some cases it is enough to check whether the expert is respected in his particular field, as we are safe in accepting the findings of a capable expert. In other, harder cases, knowing a particular expert is capable is not sufficient and one might want other experts in the field to test using the same data so that it can be checked whether the results are consistent. Unequivocal results cannot always be expected, as experts can of course also disagree (as was the case in the British Sally Clark case and in the Dutch Lucia de B. case, in which there was (and is) disagreement between statistical experts on the calculations made).

Witness testimony and expert testimony are two important types of testimonial evidence. In the law, there are other subtypes of testimonial evidence which have slightly deviant features from these two main types. For example, police officer's testimonies are sometimes seen as more credible than normal witness testimonies. Statements by the suspect, who is essentially also a witness, can also be regarded as different from normal witness testimonies in that a suspect has more interest in the case than a normal witness. These other types of testimonial evidence will not be further discussed here; for current purposes it is enough to recognize that for any testimony the elements of veracity, objectivity and observational sensitivity can and should be analysed and that in the case of an expert testimony it is important to determine what other experts in the field say.

In addition to testimonial data, Anderson and colleagues (2005, pp. 64–65) distinguish *tangible evidence*, which includes all non-testimonial sources such as legal and non-legal documents, objects and images. An important kind of tangible evidence are *documents*. Testimonies and descriptions of objects but also decisions by investigators or judges can all be recorded in documents,[5] and if at any time an investigator or decision maker wants to check the findings or decisions made in a previous stage he will consult these documents. In addition to documents, there are other forms of tangible evidence, such as videos (e.g. from CCTV cameras) and sound tapes (e.g. a taped testimony by a witness); all kinds of signs and traces such as finger- and shoeprints; material originating from the human body such as blood, hairs and sperm; objects such as bullets, gloves, cars and so on. This non-documentary tangible evidence often needs to be interpreted before it tells us

[5] In fact, the Dutch criminal procedure relies almost solely on documents.

anything meaningful. For example, finding the suspect's hair at the scene of the crime does not automatically mean that the suspect was at the scene when the crime was committed. Similarly, a bloody knife shown in court does not have to be the knife with which the victim was murdered.

The most important feature of tangible evidence is its authenticity. The evidence may have been fabricated in order to purposely mislead; for example, evidence may be planted at the crime (as happened with the glove in the O.J. Simpson) or a document may be forged. There may also be unintentional errors in the processing of evidence: a blood sample may be labelled in the wrong way or a DNA sample may be contaminated with other DNA. Here it is important to consider the so-called "chain of custody ", the documents showing what happened to the evidential data from the moment it was gathered to the moment it was used in court.

This section has aimed to give a general introduction to what is meant by evidence in a criminal legal context. The notions of direct and indirect support by evidence will be further discussed in Section 3.1.1. Various types of evidence and the way in which they can be called into question are further discussed in Section 5.2.1.

### *2.1.3 General Knowledge and Generalizations*

Evidential data is the most important source of knowledge in evidential reasoning; the knowledge gained from this data provides the "ground" on which all other reasoning about a case can be built. However, reasoning with evidence also involves reasoning with *general knowledge* or *knowledge from experience*, that is, knowledge which is not based on evidential data. General knowledge is widely accepted in a certain community. For example, it is general knowledge that elephants are larger than mice, that the date of Christmas is the 25th of December and that in most shops in the Netherlands you can pay with a bankcard. Experience-based knowledge is, as the name suggests, gained through firsthand experience. From my experiences, I know that one should watch out for bicycles in Dutch traffic and that in some small shops in the Kanaalstraat in Utrecht you have to pay cash. Note that it is possible for general knowledge and experience-based knowledge to overlap; for example, I also know from experience that in many Dutch shops you can pay with a bankcard. For general and experience-based knowledge it is usually hard to determine a specific source. This is also not always necessary: in most jurisdictions, there is something like "judicial notice" or facts which do not need to be proved because they are generally known or matters of common knowledge (cf. Rule 201 FRE). The Dutch Code of Criminal Proceedings also determines that facts and circumstances which can be considered general knowledge need not be supported by evidence (article 339 paragraph 2 DCCP).

While an exact source of general knowledge can often not be given, it is generally assumed that such information is stored in what is metaphorically called a *stock of knowledge*, general commonsense knowledge about the world that is "stocked" in the reasoner's mind. This world knowledge can take the form of simple facts (e.g. Amsterdam is the capital of the Netherlands) but also of *generalizations*.[6] Generalizations are generalized statements about how we think the world around us works, about human actions and intentions, about the environment and about the interaction between humans and their environment (Cohen, 1977, pp. 274–276). They can be based on empirical research but they can also be drawn from everyday experience or general knowledge. Examples of generalizations are "the forceful impact of a hammer can cause a person's skull to break", "witnesses under oath usually speak the truth", "people normally do not lie", "a sane person would not hit himself in the head with a hammer", "Henry is usually at work before ten o'clock" and "people from Suriname are more prone to becoming involved in crime than autochthonous Dutch people". Generalizations are almost never universally true and there are often exceptions to the generalization. For example, witnesses under oath do not always make true statements, they can lie or misremember; and a plastic toy hammer will not cause someone's skull to break. Because they are not universally true, generalizations are often qualified with terms such as *usually*, *often* and *sometimes*; in Schum's terms (1994, pp. 81–82), they are "hedged". In this book, generalizations will often not be qualified with an exact term or indication of probability; rather it will be assumed that generalizations only hold defeasibly (see Sections 2.3 and 5.1). Finally, generalizations such as the ones given above can be rewritten as conditional rules of the form "if . . . then . . ."; for example: "*if* there is a forceful impact *then* this can cause a person's skull to break" but also "*if* person *w* is a witness under oath and *w* says that event *e* happened *then* usually *e* will have happened". In this sense, generalizations can be likened to Toulmin's (1958, 2003) *warrants*, Walton's (1996) *argumentation schemes* and Reiter's (1980) *defaults*.[7]

Anderson and colleagues (2005, p. 102) mention several features of generalizations. Generalizations can vary in their *specificity*, that is, they can range from general to specific, where the general ones usually express context-independent knowledge and beliefs and the specific generalizations represent more detailed domain knowledge or context-dependent beliefs. For example, the generalization "if a person perceives *p* with his senses, then (plausibly) *p* is the case " expresses a general cognitive principle, the generalization "if a young person with good eyesight sees *p* then (plausibly) *p*" expresses more detailed knowledge and the generalization "if a person with normal eyesight and under normal conditions, who was standing two meters from the doorway of this particular apartment building and looking at the doorway, perceives that a screaming woman with red hair came running out of the doorway, then (plausibly) a screaming woman with red hair came running out of the

[6] The term "generalizations" is also used by Anderson, Schum and Twining (e.g. Anderson, 1999; Anderson et al., 2005; Twining, 1999). Wagenaar and colleagues (1993) use the term "commonsense rules".

[7] In the rest of this thesis, the term "generalization" is used for commonsense statements of both a conditional and non-conditional form, keeping in mind that it is in theory always possible to rewrite a generalization as a conditional rule (see also Section 5.1.).

doorway" expresses very specific context-dependent knowledge. Anderson and colleagues (2005, pp. 266–267)also discuss *case-specific generalizations*: generalized descriptions about personal habits or character, or local practices. Examples of such case-specific generalizations are "the employer-defendant in this case regularly discriminated against women in its employment practices", "the defendant in this case usually kept sleeping pills in the cupboard in her bedroom" or "Henry is usually at work before ten o'clock".

Generalizations also vary in their *reliability* ranging from scientifically well-established generalizations (e.g. "oil is usually of a lesser density than water") and generally accepted knowledge (e.g. "in most shops in the Netherlands you can pay with a bankcard") to biases or prejudices which are held even though no data has been supplied for them (e.g. "men are better drivers than women"). Another feature of generalizations is their *commonality*, which is governed by the number of people that share the knowledge or belief; this can vary from generalizations that are shared only by a small number of people, for example complex scientifical generalizations or certain beliefs that only a small part of the world population have, to generalizations that are widely and cross-culturally accepted. In this respect, Cohen (1977) argues that within a group of people (e.g. police investigators, jurors, judges) there is a *cognitive consensus* about the generalizations in our stock of knowledge.As long as the relevant group of people is culturally homogenous, there should be no disagreements about the generalizations themselves because they are an essential part of the culture of such a group. Rather the disagreements will be on whether a certain generalization is applicable in the given circumstances.

According to Twining (1999, see also Anderson et al., 2005; Twining, 2006), generalizations expressing general knowledge are "necessary but dangerous": they are necessary when reasoning with evidence and commonsense knowledge, but they are also dangerous in that they can be based on prejudice or false beliefs and they may not be shared by everyone within a particular community. It is argued (Anderson et al., 2005, pp. 273–276) that Cohen's idea of a cognitive consensus about our stock of knowledge is somewhat problematic. In any given group of people there will always be disagreements about which knowledge to accept and a general consensus is almost impossible in a dynamic, multicultural and multi-class society. They also argue that the idea of a stock of knowledge as a neatly organized database of generalizations is wrong. Rather, our stock of knowledge is "a complex soup of more or less well grounded information, sophisticated models, anecdotal memories, impressions, stories, myths, proverbs, wishes, stereotypes, speculations and prejudices " (Twining, 1999, p. 91). This complex structure makes it hard to separate fact from fiction and empirical generalizations from value judgements. Generalizations are often phrased in an indeterminate, ambiguous or value-laden way (see Anderson et al., 2005, pp. 276–277). For example, a generalization may be indeterminate in its universality ("witnesses sometimes/usually/always speak the truth") or phrased in value-laden terms (e.g. "those provincial yokels from the University of Groningen are usually not very good at legal reasoning"). Furthermore, a generalization's source can be indeterminate; for example, is the above generalization about people from Suriname based on research or on prejudice? Anderson (1999) warns against the term “empirical generalizations”, as this implies that generalizations have been empirically tested which is most often not the case.

In sum, generalizations and the concept of a clear cognitive consensus about these generalizations can be dangerous. Generalizations may be invalid as they can be based on dubious ideas about the world. That there is a clearly defined cognitive consensus about our stock of knowledge is also questionable. Furthermore, this stock of knowledge may also contain other, more complex knowledge structures such as stories. One of the main dangers here is that generalizations (or the contents of the stock of knowledge) are often left *implicit* in the reasoning process. Therefore, the validity of a generalization or any exceptions in the case at hand cannot be checked. Anderson and colleagues (2005, pp. 279–282) provide protocols for assessing the plausibility of generalizations and stories. Similarly, Crombag et al. *Anchoring Process* (1994; Wagenaar et al., 1993, see Sections 3.2.1 and 7.1.2) is aimed making generalizations explicit and specifying them as well as possible. The analysis of generalizations and the reasoner’s stock of knowledge in general play a major role in this book; reasoning with general knowledge and reaching a cognitive consensus about this knowledge through a rational discussion is one of the main uses of the hybrid theory as proposed in Chapters 4 and 5. Specific Sections on this general knowledge are 3.1.3, in which the analysis of generalizations using arguments is discussed, Section 5.1 in which a brief logical account of reasoning about the validity of and exceptions to generalizations is given, and Section 3.2.2, which introduces the concept of *story schemes* as more complex knowledge structures that are part of the stock of knowledge.

## 2.2 The Process of Proof: Discovery, Pursuit and Justification

The process of proof involves discovering, testing and justifying hypotheses about what happened in a particular case. The process usually starts when some initial observations are made, that is, when some initial evidence is found. On the basis of this preliminary evidence one or more initial hypotheses will have to be imagined. Some of these initial hypotheses will be immediately discarded as implausible, while others have to be tested by searching for further evidence and then determining which of them are compatible with the new evidence. When a particular hypothesis has been chosen as the most likely, this choice should be justified by explicitly showing that it is most compatible with the evidence.

As an example of this process, say that we are faced with a dead body of a man and we hypothesize that either the man had an accident, that he killed himself or that he was killed by another person or animal; we could also hypothesize that aliens killed him, but this scenario will usually be discarded because of its intrinsic implausibility. Looking for further evidence, we might find that the man died of an overdose of sleeping pills. This evidence is arguably incompatible with the “accident” scenario, because people usually do not take large quantities of sleeping pills by accident. The hypothetical suicide scenario is now the most likely. However, it is still possible that the man was drugged by another person. Further evidence like a suicide note or testimonies by the man's psychiatrist that he had been suicidal for a long time could settle this choice between hypotheses. If we settle for the suicide scenario, we could motivate this choice by saying that the note combined with the sleeping pills clearly point to a case of suicide.

In the philosophy of science, the contexts of *discovery* and *justification* are often distinguished. This distinction is commonly attributed to the German epistemologist Reichenbach (1978), who used the two contexts mainly to distinguish between the logical and psychological aspects of science: justification is about objective and logical relations between a body of evidence and a theory and discovery is about the subjective and psychological process of finding this theory and the relations between the theory and the evidence. It is only the justification context that the philosopher and the logician are interested in; analysing the process of discovery is the domain of psychologists. Later philosophers such as Hanson (1962) distinguished between three contexts: *discovery*, *pursuit* and *justification*. It is possible to characterize the process of evidential reasoning along these three contexts: discovery concerns the formation of various hypotheses,[8] pursuit concerns testing the various hypotheses and choosing the most likely one, and justification is about arguing why this particular hypothesis should be accepted. The outcome of each part of the process serves as the input to the next part. The hypotheses found during discovery serve as input for the testing phase, where the hypotheses should be proved or disproved and where ultimately the best hypothesis should be chosen. During this "pursuit" of the best hypothesis connections between the evidence and the hypotheses are made and various arguments for and against the different hypotheses are put forth. When made explicit, such connections and arguments can be used in the next stage of the process to justify why a particular hypothesis should be accepted.

Table 2.1 shows the process of moving through the different contexts. This table has been adapted from Nijboer and Sennef (1999), who use different terms for the elements of the various contexts. In the context of discovery, we start with the *explananda* and end up with a number of possible hypotheses, the *explanans*, which explain these initial observations. These hypotheses serve as the input for the next part of the process (the context of pursuit), where they can be viewed as *probanda*, propositions that have to be tested and proven. Making a choice between the various available hypotheses in a case provides us with a particular hypothesis that can be considered proven, the *probans*. This probans and the reasons for choosing it together make up a *proof*. In other words, a proof is the chosen hypothesis (which is deemed most likely) and a justification of why the particular hypothesis was chosen.[9] In the context of pursuit, we are engaged in the process of trying to

---

[8] Here I use the general term "hypothesis". Such a hypothesis can have various forms (e.g. a hypothetical scenario, the identity of a person's whose blood was found and so on). Below the specific process of proof for crime investigation will be briefly discussed.

[9] Wigmore characterizes a proof as "the persuasive operation of the total mass of evidentiary facts, as to a probandum" (Wigmore, 1931, p. 9). The important parallel with the current conception of proof is that the total mass of evidence that pertains to a hypothesis is considered in a proof.

**Table 2.1** The process of proof(adapted from Nijboer and Sennef, 1999)

| context of discovery | context of pursuit | context of justification |
|---|---|---|
| *Explananda* (observations that need to be explained) | | |
| ↓ | | |
| *Explanans* (hypotheses that explain the observations) → | *Probanda* (elements of a hypothesis that need to be proven) | |
| | ↓ | |
| | *Probans* (elements of a hypothesis considered proven) → | *Proof* (proof that needs to be explicated and justified) |
| | | ↓ |
| | | *Justification* (Justification of the proof) |

justify (proving) a hypothesis, and in the context of justification we make this justification (proof) explicit. The process of proof is not only focused on rationally justifying one's reasons for choosing a hypothesis but also on the clarification of these reasons. In this respect, Nijboer and Sennef (1999) speak of *explanatory justification*.According to this purpose of justification, a proper justification should not only meet some standard of rationality but it should also provide a clear explication of one's reasons for the choice of hypothesis which makes sense to not just the reasoner but also to third parties.

In the process of proof it is common to go back some steps. If, for example, no proof can be found for a hypothesis in the context of pursuit, the discovery phase starts anew so that a new hypothesis can be found. The overarching process of proof also consists of many "sub-processes", in which small hypotheses which are part of a larger hypothesis are individually tested and subsequently incorporated into the main hypothesis. Finally, the various contexts also overlap as one context provides the input for the next. Because of this complex structure of the process of proof, there is no definite consensus in the literature on exactly what constitutes each context. For example, Thagard (1988) argues that discovery is purely the construction of new hypotheses and that testing hypotheses is part of the context of justification, whilst others (e.g. Josephson and Josephson, 1994; Walton, 2001) seem to argue that discovery includes both the formation and the testing of new hypotheses. It is for present purposes not important what exactly falls under each context, as long as it is recognized that reasoning with evidence can be characterized as an iterative process of discovering, testing ("pursuing") and justifying hypotheses.

### *2.2.1 Investigation, Trial and the Process of Proof*

Just as evidence is not restricted to the (criminal) legal domain, the process of proof consisting of the contexts of discovery, pursuit and justification is not exclusive to reasoning with evidence in a criminal legal context. However, some specifics of the criminal process of proof can be indicated. In a criminal context, the process of proof starts with some initial clues that point to the fact that something out of the ordinary happened; this will usually be some evidence that points to the possibility of a crime having been committed. If on the basis of this preliminary evidence it is decided that an investigation will be started, the investigators have to imagine some initial hypotheses to try and reconstruct "what happened" in the case. In crime analysis, these hypotheses take the form of *hypothetical scenarios*[10] describing what might have happened, that is, the crime and the circumstances surrounding the crime (De Poot et al., 2004; Kerstholt and Eikelboom, 2007).

De Poot and colleagues (2004) give an interesting model of crime investigation, in which they divide cases into four categories. In *clear-cut cases* the police catch a criminal red-handed or the criminal turns himself in. In *verification cases*, a possible scenario as well as the identity of the suspect are known at the start of the case. An example of such a case is when the victim or a witness reports a crime by someone he or she knows or when the police find a suspicious person near the crime. In *investigation cases* a possible scenario is known at the start of the case but the suspect has to be tracked down. An example is a case where a witness reports a crime committed by a person the witness does not know. Finally, in *search cases* neither a scenario nor a suspect are known at the start of the case. An example is a case where only physical evidence is found, for example, a successful burglary. During the investigation the case changes into the various types. If, for example, a suspect is found in an investigation case, the case changes into a verification case. It can be seen how this characterization of the process of proof in crime investigation fits the contexts of discovery, pursuit and justification. In a search case hypothetical scenarios are discovered, which are then investigated and verified in the context of pursuit. Ideally a clear-cut proof is then given for a particular scenario so that the investigation can be closed and the trial phase can start.

The above discussion only concerns criminal investigation and does not take the decision-making phase in a legal case into account. Once a case ends up in court, we have essentially arrived in a specific context of pursuit in which the probanda that have to be proved are fixed in advance. In American law, these probanda are the material facts and can be viewed as the conditions of the legal rule as given in substantive law. In Dutch law, the ultimate probandum is given by the indictment and

[10]In the overarching process of proof in crime investigation, hypotheses can be equated with scenarios. However, other hypotheses are also important in any of the "sub-processes of proof". For example, a hypothesis can also concern the identity of the possible perpetrator or the exact time of death and each of these individual hypotheses needs to be tested before they can be incorporated in the main scenario.

the penultimate probanda are the various components of this indictment.[11] While a judge or a jury may have some possibilities for performing new investigative actions, usually the majority of these investigative actions (e.g. discovering new hypotheses, gathering evidence) will have been done in the pre-trial stage. In this way, the total criminal and legal process from first clues to conviction or acquittal is a process of proof (Nijboer, 2000, p. 61), where the pre-trial phase is the context of discovery and the trial and deliberation is the context of pursuit. Because American juries do not have to explicitly state reasons for their ultimate decision, Nijboer argues that there is no context of justification in the American legal process. Dutch judges, however, are bound by law to state their reasons for a conviction and the Dutch legal process therefore ends with the context of justification.

## 2.3 Reasoning in the Process of Proof

In the process of discovering, pursuing and justifying hypotheses, various types of reasoning are employed, which are discussed in this section. Almost all reasoning with criminal evidence can be characterized as *defeasible reasoning*.[12] In this kind of reasoning, new information can cast doubt on propositions previously taken to be true. This reasoning is also known as *nonmonotonic reasoning*, which can be contrasted with monotonic reasoning in which a new piece of knowledge cannot reduce the set of what is taken to be true. In defeasible reasoning, the main interest is whether we are currently justified in believing some proposition and which new knowledge might change this belief. One influential idea that has been put forward in the context of defeasible reasoning is that we are justified in believing that a proposition is true if the proposition can stand against criticism in a dialectical inquiry.[13] For example, Rescher (1977) says that claims in scientific reasoning, which have often been established through inductive reasoning, can be accepted if they cannot be successfully challenged in a proper scientific dispute. The importance of this dialogical setting in reasoning with evidence, arguments and stories will be further discussed in the rest of this book.

Like Rescher, Anderson and colleagues (2005, p. 100) equate defeasible reasoning with *inductive reasoning*, which they contrast with deductive reasoning and abductive reasoning (see below); these three types of reasoning form the "triangle" well-known in philosophy.[14] However, inductive and abductive reasoning can also

[11] Dutch law (art. 338 and 350 DCCP) says that the judge should determine whether the suspect committed the acts as given in the indictment.

[12] The term "defeasibility" seems to have been introduced in legal philosophy by Hart (cf. Loui, 1995), who argued that a legal concept (e.g. "contract") not only encompasses the positive conditions which have to be met (e.g. there has to be an offer and acceptance of the offer) but also that which can defeat the claim that we are dealing with the appropriate legal concept (e.g. the parties hold a mistaken belief of the facts).

[13] Van Eemeren and colleagues (1996) characterize dialectic as "the art of arguing for and against".

[14] The triangle is often attributed to Peirce, see (Burch, 2008) and (Peirce, 1931).

be seen as simply being subtypes of defeasible reasoning. In the rest of this book I will therefore consider inductive reasoning and abductive reasoning in their more narrow sense, that is, as reasoning from particular examples to a universal statement and reasoning from observations to hypotheses respectively. The terms defeasible, nonmonotonic or presumptive reasoning (cf. Walton, 2001) will be used to cover all kinds of non-deductive reasoning.

### *2.3.1 Abductive Reasoning and Inference to the Best Explanation*

Reichenbach argued that the process of discovery is something for psychologists to analyse and that philosophers and logicians should only concern themselves with justification. Consequently, he argued, there is no such thing as the "logic of discovery ". However, other authors (e.g. Hanson, 1962; Schum, 1994, Chapter 9; Thagard, 1988, Chapter 4) have argued that Peirce's (1931) notion of *abductive reasoning* has certain features that could mark it as an imaginative type of reasoning that allows us to discover new hypotheses or theories. The best way to describe the basic idea of abductive inference is to contrast it with deduction and induction. The following example is cited in several texts on abduction (e.g. Burch, 2008; Walton, 2001).

> Deductive Reasoning: Suppose a bag contains only red marbles, and you take out a particular random sample. You may infer by deductive reasoning that all the marbles in the sample will be red.
>
> Inductive Reasoning: Suppose you do not know the colour of the marbles in the bag, and you take out (a sufficiently large) random sample and all the marbles in the sample are red. You may infer by inductive reasoning that all the marbles in the bag are red.
>
> Abductive Reasoning: Suppose you see some red marbles in the vicinity of a bag of red marbles. You may infer by abductive reasoning that the marbles are from the bag.

In this example, the difference between the three forms of inference is quite obvious: here deductive inference is strict reasoning from a universally quantified premise to a conclusion and inductive inference is defeasible reasoning from a particular example to some universally quantified conclusion, abductive inference is defeasible reasoning from some observation to a possible hypothesis. Schematically, abductive reasoning can be represented as follows (adapted from Josephson and Josephson, 1994):

> D is a collection of data (facts, observations, givens)
> Hypothesis H explains D (would, if true, explain D)
> ---
> Therefore, H is probably the case

In this scheme it can be seen why abduction can be characterized as *explanatory reasoning*, where one observes some states of affairs and tries to explain how these could have followed from some hypothesized events or states of affairs. For example, we can say that the man's death has been caused by him taking the sleeping pills. In other cases, however, the exact causal relation will not be that clear. In the above example of the red marbles, exactly what caused the marbles to be next to the bag is not clear, but we still might say that the hypothesis that the marbles are from the bag explains the marbles being in the vicinity of the bag. The hypothesis could be further specified by saying that the marbles are from the bag and that the bag falling over caused the marbles to roll out of the bag.

An important feature of explanatory abduction is that it provides us with a *new* hypothesis. In this sense, abductive reasoning has variously been called imaginative, creative and ignorance avoiding.[15] Eco (in Eco and Sebeok, 1983, Chapter 10) argues that there are varying degrees of creativity in abductive reasoning. In the least creative form, one draws upon one's previous experiences with a particular situation to infer a hypothesis of how this situation came to be. In its most creative form, abduction concerns the formation of truly new hypotheses which are not based on any previous experience. Schank (1986) also argues that there are essentially three ways of explaining observations. The first way is to look at similar cases; if we want to know why the Dutch football team lost to Russia, we take another match or other matches where they lost and we try to find similarities between the last match and these other matches (e.g. "in all the matches, a key defender was injured"). The second way to explain an event is to find a more general explanation pattern for such an event. In the case of the Dutch football team, we could ask why sports teams, even if they seem better, sometimes lose. Maybe it is overconfidence? The third, and arguably the hardest way to explain an event happens when there is no particular case or pattern that fits the current case; we then have to build a new explanation from scratch. In the football example, this would involve analysing everything that happened during the match and thus form an explanation about why the Netherlands lost.

Once a hypothesis has been inferred through abductive reasoning, the hypothesis should be as Schum (2001) calls it, "put to work": assuming this new hypothesis, we should try to predict what else may be observed if the hypothesis were true. Here we perform *predictive reasoning*, where one makes an observation or assumes a hypothesis and tries to predict what will be the case as a consequence of this observation or hypothesis. Note that explanatory and predictive reasoning are closely related: a hypothesis *explains* some observations if the hypothesis correctly *predicts* these observations. We can try to predict more than just the observations which originally led to the hypothesis and thus try to predict new observables. For example, if we assume that the man from the example at the beginning of Section 2.2 committed suicide, we should be able to find a suicide note. Thus such a prediction can guide the search for new evidence. So the context of discovery concerns not just the discovery of new hypotheses given some evidence, but also the discovery of new evidence given some hypothesis.

Walton (2001) argues that abductive inference can be characterized as an "intelligent guess". However, when a hypothesis thus formed is thoroughly tested, such an intelligent guess can become a plausible explanation for some observed phenomenon. By predicting potential observables and trying to find evidence for them,

---

[15]This last term is by Gabbay and Woods (2006), who also contrast it with deductive reasoning, which they call truth-preserving and induction, which they call likelihood-enhancing.

we have in a sense started the context of pursuit, in that we look if the hypothesis we are currently pursuing explains any more observations. In this context, we should also look if and how possible new evidence can be reconciled with the current hypothesis and, perhaps most importantly, we should look if there are other hypotheses that explain the evidence better than the current one. For example, if we find out an urn with red marbles is also in the vicinity, the marbles may just as well have come from the urn instead of from the bag.

Testing and comparing hypotheses has been called *inference to the best explanation* (IBE)[16] in the literature (Thagard, 1988). IBE should not be seen as a single inference step, but rather as a general term for a combination of several ways of reasoning with and about hypotheses. Essentially, IBE covers all the important types of inference in the process of proof: hypotheses are constructed, these hypotheses are tested and compared with each other and ultimately the best one is chosen. Exactly how hypotheses should be compared and when a particular hypothesis should be chosen as the best explanation for the observations is a subject of much debate. Both Thagard (1988); Josephson and Josephson (1994) argue that in order to be considered as the best, a hypothesis should better conform to the evidence in a case, that is, the hypothesis should explain the observations in the case better than the other available hypotheses. Other considerations when determining the best hypothesis are, for example, whether the hypothesis itself is inherently plausible, how many plausible alternatives there are and how thorough the search for alternative hypotheses has been. In Sections 3.2.4 and 4.4, various ways of testing and criteria for choosing an explanation will be further discussed.

### *2.3.2 Causal Reasoning with Evidence*

Reasoning with causal information is an important aspect in all stages of reasoning about evidence. When investigators find certain clues, they try to figure out what caused these observations (explanation) and what else could have been caused by the hypothesis for these observations (prediction). Also, a proof is often constructed by saying that an observed fact (the evidence) holds since something else (the crime) happened which caused it.

When reasoning with causality, it is possible to distinguish between *causal* and *evidential reasoning* (Pearl, 1988a).[17] Causal reasoning is meant as reasoning with causal generalizations of the form "*c* is a cause for *e*" and evidential reasoning stands for reasoning with evidential generalizations of the form "*e* is evidence for *c*". Causal and evidential reasoning are closely related: if we have a causal

[16] I do not consider abduction and IBE to be one and the same. Abduction involves solely the creation of new hypotheses, whereas IBE is also about testing and comparing hypotheses.

[17] Note that Poole, Mackworth and Goebel's (1997) distinction between causal and evidential reasoning is different from the one presented here. Their evidential reasoning is essentially what is here called explanatory reasoning and their causal reasoning is what is here called predictive reasoning.

generalization "*c* causes *e*" then we will usually also accept that "*e* is evidence for *c*".[18] For example, fire can cause visible smoke so the observation of smoke can be seen as evidence for the fact that there is a fire. On this subject, Thagard (2004) argues that by accepting a testimony $E^*$ as evidence for an event $E$ we implicitly make a causal judgment that the witness tells us $E^*$ because that witness really believes $E$ happened. Similalry, Pardo and Allen (2007) argue that "explanations are "self-evidencing" in the sense that what is explained (the evidence) provides a reason for believing that the explanation is correct". Note that both forms of reasoning are defeasible: fire does not necessarily cause (visible) smoke as some types of fire (e.g. a gas flame) do not cause smoke; similarly, observing smoke does not necessarily imply that there is a fire as there are other causes of smoke (e.g. a smoke machine). The close relation between causal and evidential reasoning is also evident in the *explains* relation, which was already presented in the abductive reasoning scheme in Section 2.3.1. Thagard (2004, p. 237) argues that some proposition *c* explains some other proposition *e* if the states of affairs described by *c* are part of the "causal process" that produces the states of affairs described by *e*. Thagard then quotes Quine and Ulian (1970), who argue that there is mutual reinforcement between an explanation and what it explains. A proposition *e* that we believe to be true is more credible if there is a plausible explanation *c* for it and at the same time an explanation *c* is more credible if it explains a proposition *e* we believe to be true. Furthermore, Simon (2001) has argued that "... [people] defy the syntactic rules of unidirectional inference ...", i.e. that causal and evidential reasoning are both used in conjunction.

In formal AI models, predictive reasoning is almost always modelled as reasoning with causal generalizations. Explanatory reasoning, however, can be modelled in two ways. With evidential generalizations explanatory reasoning can be modus-ponens style: if the antecedent (the effect, e.g. "smoke") is known, the consequent (the cause, e.g. "fire") can be inferred. When thus alternative explanations can be derived, a choice should be made with some priority mechanism. With causal generalizations prediction can also be modus-ponens-style but explanation must be done by "affirming the consequent": given the consequent (the effect) the antecedent (the cause) is inferred since if the cause is true it implies the effect by modus ponens on the causal generalization (cf. the relation between explanation and prediction, Section 2.3.1). Of course, alternative explanations may be found in the same way and the best explanation must be determined with some priority mechanism on the various inferred causes that act as explanations for the effects. In Fig. 2.1, these two different ways of modelling explanatory reasoning (i.e. causal and evidential) have been represented.

$$\frac{\mathit{cause} \longrightarrow \mathit{effect}, \quad \mathit{effect}}{\mathit{cause}} \qquad \frac{\mathit{effect} \longrightarrow \mathit{cause}, \quad \mathit{effect}}{\mathit{cause}}$$

**Fig. 2.1** Different ways of modeling explanatory reasoning

[18] It can be argued that this is only the case when *c* is a typical cause of *e*.

There is a third way of modelling explanatory reasoning, namely with relations which can be expressed as an explanatory generalization of the form "*a* explains *b*". Because this explains relation is symmetrical, the knowledge about smoke and fire can essentially be modelled as either "smoke *explains* fire" or "fire *explains* smoke". This way of modelling explanatory reasoning is mainly used by Thagard (1989, see Section 7.3.1) and Josephson (2002, see Section 7.3.2).

Explanatory reasoning with causal generalizations (i.e. in the way shown on the left of Fig. 2.1) is often equated with abduction (Shanahan, 1989; Console et al., 1991). If the term abduction is interpreted in a broad sense, this need not be the case, as then it simply stands for reasoning from observations to a hypothesis and this reasoning can also be modelled with evidential generalizations or explanatory generalizations. However, in AI explanatory reasoning is often modelled as reasoning with causal generalizations (see Shanahan, 1989) and the term "abduction" is therefore equated with reasoning by affirming the consequent of a causal generalization. Walton (2001) also argues that abduction can be seen as going backwards from a given conclusion to search for the premises that conclusion was based on (from Latin *ab* and *duco,* leading back). Hence, from here on the term abduction will be interpreted narrowly in that it stands for explanatory reasoning with causal generalizations; reasoning from observations to a hypothesis (in whichever way) will simply be called explanatory reasoning. In AI research reasoning from premises to a conclusion (i.e. predictive reasoning with causal generalizations or explanatory reasoning with evidential generalizations) has been characterized as "deduction" (Shanahan, 1989). However, this type of reasoning is not always deductively valid: fire does not necessarily cause (visible) smoke and observing smoke is not necessarily evidence for a fire. It is for this reason that such reasoning from premises to a conclusion is called here not called "deduction" but rather "modus-ponens style reasoning".

### *2.3.3 Reasoning About Motives and Actions*

When reasoning about causality, we can differentiate between causality as regarded from the *physical* as well as the *intentional stance* (Dennett, 1978). Crimes involve rational persons and hence we need to see events not simply as the result of the operation of physical causal laws but also as the result of choices made by the persons; hence it has to be determined what kind of motives and goals could have caused the behaviour of the (suspected) criminal. Note that here "rational" does not mean that the persons involved in crimes always make conscious rational decisions but rather that these persons' actions are guided by some sort of (perhaps irrational) motive or goal.

When discussing reasoning about motives and actions, it is useful to distinguish the separate concepts of *motive* and *goal*. Motives in a broad sense are a combination of a person's values and emotions combined with some external motivating states and events. For example, the fact a robber has no money together with his urge to satisfy his drug addiction is a motive for him to rob the supermarket. In a more narrow sense, the motives are essentially abstract concepts, principles which a person or a group of persons hold. Persons can be expected to, actively or passively, promote the motives which they find important. Examples of motives are “wealth”, “love” and “honesty”, but also urges like satisfying an addiction or seeking revenge.[19]

A motive (in any sense) can lead to different goals which can be satisfied in multiple ways. For example, the motive “money” can be satisfied by robbing a supermarket but also by finding a job. Walton and Schafer (2006) do not make the distinction between motives and goals when they claim that motives are immediate goals to which an agent is strongly committed. They follow Wigmore (1931), who defined a motive as “a specific emotion or passion that is likely to lead to a specific act”. Motives (in the broad sense) differ from goals in that motives (in the broad sense) involve some factors external to the agent. The difference between motives (in the narrow sense) and goals is much less clear; for now, we will assume that goals are actively and immediately pursued whilst motives (in the narrow sense) are more abstract principles an agent adheres to that need not always be actively promoted by the agent.

A concept related to motive is that of *character*. As Walton and Schafer note, character evidence is often inadmissible at trial. However, it may be admissible for the purpose of proving someone’s motive. Statements about someone’s character are often expressed as case-specific generalizations about specific persons. For example, “John is the kind of person who would never rob a supermarket” or “John would never kill for money”. Bex et al. (2009) model character as an *ordering* of a person’s motives; for example, we can say that John is the kind of person who prefers the satisfaction of his drug-addiction over honesty, or that he does not prefer money over the life of another person.

Motives play an important role in all stages of the process of proof. In the discovery stage, an alleged motive can play an important part in constructing provisional hypotheses. In the example from the beginning of Section 2.2, scenarios for accident, suicide and murder could be constructed: during this construction, various motives for murder or suicide can be considered. If, for example, the victim has no motive whatsoever for committing suicide this hypothesis need for the moment not be considered any further. If a plausible motive for murder is then found (e.g. “money” because the man was exceptionally rich), the hypotheses can be fleshed out by searching for persons who would financially benefit from his death. If such a person (for example, the victim’s wife who would inherit most of his fortune), is found it then has to be determined if this is a correct motive given the suspect’s character. For example, is the victim’s wife the kind of person who is motivated by monetary gain to kill her husband. Finally, in order to correctly justify the reasoning it also has to be considered that the suspect could have satisfied her motives by acting differently. For example, it might be argued that it would have been easier for the suspect to divorce her husband and take half of his fortune with her.

---

[19] In Bex et al., 2009, motives were equated with values (cf. Bench-Capon, 2003), which are also abstract principles that can be promoted or demoted.

In sum, reasoning about the motives and goals for actions plays an important part in the process of proof. It is a subject which deserves attention, particularly because often there is no direct evidence for the suspect's motives and goals: in the end, only the perpetrator knows exactly why he did it. Note that the idea of motiveless crimes can essentially be captured with the current broad definition of a motive and character: a person who stabs another person to death just because he feels like it also has a "motive", namely that he felt like stabbing someone at that time. This is arguably a strange or incredible motive, but if it can be made plausible that the perpetrator is the kind of person who ranks his morbid urge to kill someone higher than the life of another person.

# Chapter 3
# Two Approaches to Reasoning with Evidence: Arguments and Stories

In Chapter 1 it was argued that in the current research on the analysis of reasoning with evidence, essentially two trends can be distinguished. The first method for analysing evidence, which has its roots in Wigmore's (1931) evidence charts and has been mainly developed by the New Evidence Theorists (Anderson et al., 2005; Tillers, 2005), focuses on *arguments* from evidence to an ultimate probandum. These arguments can be used to structure and analyse the reasoning about the evidence in a case. The second method for the analysis of evidence, which was introduced by Bennett and Feldman (1981) and further developed by Pennington and Hastie (1986) and Wagenaar et al. (1993), mainly uses *stories* to structure and analyse the available evidence.

Whilst the central concepts of argument and story play an (implicit) role in most of the above-mentioned research, they are not fully integrated in any of the current theories on reasoning with criminal evidence. The aim of this chapter is to clarify the separate roles of arguments and stories in the process of proof. Thus, their respective advantages and disadvantages will become clear and the exact way in which they should be integrated in a natural and rationally well-founded hybrid theory can be explored. The discussion of arguments and stories will take the form of a comparison of two separate approaches to reasoning with evidence, the argument-based approach and the story-based approach.

In the *argument-based approach*,arguments are constructed by performing consecutive reasoning steps, starting with an item of evidence and reasoning towards some conclusion. Each of these reasoning steps has an underlying *evidential generalization* of the form "*e* is evidence for *p*" that justifies the step from premises to conclusion. Hence, reasoning in this way can be characterized as *evidential reasoning*. Reasoning with arguments is *dialectical*, in that not only arguments for a particular probandum but also arguments against the probandum and other kinds of counterarguments are considered. Finally, argumentative reasoning has been called *atomistic* because the various elements of a case (i.e. hypotheses, evidential data) are considered separately and the case is not considered "as a whole".

The *story-based approach* involves constructing stories about what (might have) happened in a case that explain the evidential data. Reasoning with stories, which detail the course of events before, during and immediately after the crime, can be characterized as *causal reasoning*; the relations between the various events in a

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_3, 

story and between the story and the observations (i.e. the evidential data) can be expressed as *causal generalizations* of the form "*c* is a cause for *e*". This approach also has a dialectical component in that the different stories about the case are compared according to the amount of evidential data they explain and their internal coherence. The story-based approach has also been called *holistic* (as opposed to atomistic), because the various elements in the case (i.e. hypotheses, evidential data) are considered as a whole and the elements receive less individual attention.

Note that these two approaches are each not based on one particular author or group of authors, but abstract accounts of the two trends in research about evidence as discussed at the beginning of this section. Consequently, the following sections should not be regarded as a comparison of the work on Modified Wigmorean Analysis as done by the New Evidence Theorists and the work on (explanatory) stories and Anchored Narratives as done by Pennington and Hastie, Crombag and colleagues and Pardo and Allen. Rather, the present discussion aims to give abstract accounts of arguments and stories, so as to clearly disambiguate between them. A full comparison of related research with the later proposed hybrid theory of stories and arguments will be done in Chapter 7.

In the rest of this chapter, I will refer to a particular example case. The example is inspired by the Haaknat case from *Anchored Narratives* (Wagenaar et al., 1993, pp. 35, 71–72).[1] This case concerns a robbery of a supermarket in a town in the Netherlands. On the 21st of October, the police get a call that a supermarket has just been robbed by a man wearing a balaclava. This man threatened the owner of the supermarket with a gun and the owner gave the robber the contents of the safe. The robber took off and the owner called the police. When the police arrived at the supermarket, they were informed by a witness who saw that a suspicious character got into a red car and drove towards the park with great speed. The police drove to the park and found the red car parked outside it. They then searched the park and found a man named John Haaknat hiding in a moat full of water in the park. Because of Haaknat's suspicious behaviour, the police apprehended him and took him to the station for questioning.

## 3.1 Evidential Arguments

The basic idea of an argument is something that dates back to the philosophers of ancient Greece. Aristotle's *syllogism* is in essence an argument, where premises, one of which is a conditional statement, lead to a conclusion; his *dialectic* can be viewed as a basic process of argument and counterargument. There are multiple interpretations of the term "argument"; an argument can be, for example, a single reason for a conclusion or a chain of reasons that leads to a conclusion but also the combination of the reasons for and against a certain conclusion or a dialogue between parties trying to convince each other (cf. Wyner et al., 2008).

[1] The case as presented here is a simplified account and deviates from Wagenaar and colleagues' account on some points.

Stephen Toulmin, in his seminal work "the Uses of Argument" (Toulmin, 2003), drew attention to many of the features of everyday commonsense reasoning which, in his view, logicians had neglected. He introduced a richer structure for arguments by distinguishing separate elements such as the datum (premise), claim (conclusion), warrant (the underlying license for the inference step) and backing of the warrant. Furthermore, he argued for defeasibility of arguments by allowing them to be attacked by statements that provide an exception to the warrant. Finally, by interpreting logic as "generalized jurisprudence", Toulmin (implicitly) argued for the dialectical role of logic in a critical procedure for attacking and defending claims.

John Pollock (1987, 1995)provided the necessary formal-logical grounding to Toulmin's diagrammatic model of argument structure. Although his theory was not originally developed with the express aim of formalizing Toulmin or defeasible argumentation in general, many of his basic ideas about the formal structure and behaviour of arguments persist. Other pioneering work on formal argumentation is (Loui, 1987). Since then, the field of formal argumentation (see Prakken and Vreeswijk, 2002; Chesñevar et al., 2000 for overviews) has developed several ideas, including formal argumentation- theoretic semantics for defeasible reasoning (Dung, 1995). At the same time, the fields of argumentation theory and informal logic have also contributed to the understanding and modelling of arguments (see van Eemeren and colleagues 1996 for an overview). For example, van Eemeren and Grootendorst proposed rules for a critical discussion of arguments in their pragma-dialectical approach (van Eemeren and Grootendorst, 2004), Freeman (1991) and Reed and Rowe (2004) have written about the diagrammatic structure of argument and Walton has extensively discussed various forms of argumentation in different contexts, including legal evidence (Walton, 1996, 1998, 2002).

The basic idea of drawing inferences from premises to a conclusion was also present in Wigmore's work. In *The Principles of Judicial Proof* (Wigmore, 1931), Wigmore set out to develop a rational theory which could be used for structuring and analyzing inferences based on a mass of evidence. These chained and complex inferences can be depicted as a tree-like chart, where the leafs of the tree are the pieces of evidence and the root is a major probandum (see, for example, Anderson et al., 2005, p. 139; Bex et al., 2003; Dingley, 1999; Kadane and Schum, 1996).Wigmore regarded reasoning about evidence and proof as a kind of commonsense reasoning using argument charts. Wigmore's work on evidence and inference was further developed by Anderson et al. (2005). Following Toulmin, they introduced the idea of commonsense generalizations as inference warrants and they explicitly speak of a chart as an "argument".

The ideas by Wigmore and the New Evidence Theorists on commonsense reasoning are largely compatible with the (formal and informal) work in defeasible argumentation. For example, Feteris (1999) applies the pragma-dialectical approach to the famous Dutch ballpoint case and Bex et al. (2003) have explicitly shown how the structure of Wigmore's charts corresponds to more formal models of argumentation. The following sections can be regarded as an extension of (Bex et al., 2003).

### *3.1.1 The Structure of Evidential Arguments*

Above (Section 2.1) it was argued that through autoptic proference, nothing more than the mere existence of the evidential data can be accepted and that further reasoning with this data requires us to construct arguments from this evidential data to some conclusion. In an argument, a defeasible inference leads from premises to a conclusion; associated with a defeasible inference is a *generalization*, usually in a conditional form, which justifies the inference link between premises and conclusion. Anderson and colleagues (2005, p. 62) argue that the generalization warrants the inference and can therefore be likened to Toulmin's (2003) *warrant*. Schum (1994) calls generalizations the "glue" which holds evidential arguments together. Note that the inferences in an evidential argument are of an evidential nature: some evidence *e* and an evidential generalization "*e* is evidence for *p*" allows us to infer *p*. For example, a witness " testimony" I saw someone who looked like Haaknat got into the car on the 21st of October' and the generalization "a witness testimony that '*p*' is evidence for *p*", where *p* is some state of affairs in the world, allows us to infer that the witness saw someone who looked like Haaknat get into the car. Figure 3.1 shows how a proposition can be inferred from a piece of evidence: the evidential datum *e* (rendered as a coloured box) and the generalization *g* (rendered as a rounded box) together allow us to infer the conclusion *c*; here, the inference itself is rendered as an arrow with a closed head.

Figure 3.1 shows what can be seen as a simple argument consisting of premises (in this case the testimony), a conclusion and the generalization that justifies the inference from premises to conclusion. Note that no qualifier (i.e. usually, presumably) is given in the generalization; this does not mean that the inference should be regarded as conclusive. Note that arguments do not necessarily have to be based on evidential data and that they can also be based on assumptions. In this way, it is possible to reason from general knowledge or personal experience. In Section 3.1.3 this will be further discussed. Notice that here an evidential generalization for witness testimonies has been used. This generalization can be phrased in several ways, for example "witnesses under oath usually speak the truth", "if a witness testifies that *p* is the case then usually *p* is the case" (Bex and Prakken, 2004) and "If a witness is in a position to know whether *p* is true and the witness asserts *p* then *p* may plausibly be taken to be true" (Bex et al., 2003). The exact generalization may be

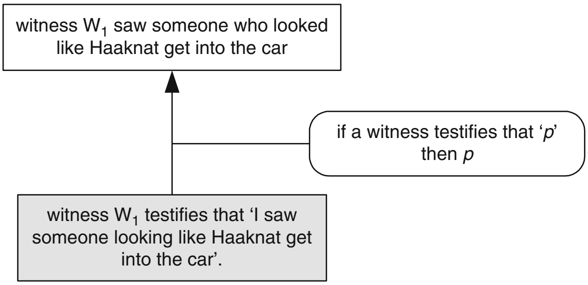



**Fig. 3.1** Inferring conclusions from evidential data

open to debate; do only witnesses under oath speak the truth? If a witness testifies to *p*, is then *p usually* the case, or perhaps sometimes or 60% of the time? What is meant by a "position to know"? However, most people would agree that conclusions can be drawn from witness testimonies. In Dutch law, a witness testimony is explicitly mentioned as a legitimate item of evidence and Nijboer (1993, p. 314) argues that the DCCP rules on types of evidence can in a sense be seen as "argumentation rules". So we are probably justified in accepting some form of the witness rule, as otherwise we would not be able to infer anything from a witness testimony. This does not mean that witness testimonies are always true; inferences are defeasible and the reasoning is dialectical, so exceptions to the general rule can be given; this will be further discussed below.

In some cases, an argument is defined simply as a pair "premises – conclusion" and the generalization justifying the reasoning step from premises to conclusion is left implicit (Fig. 3.2). This is often the case when the same generalization occurs multiple times in a complex argument; for example, when modelling a large case with multiple testimonies, the witness testimony generalization is not explicitly shown in every argument. Seemingly obvious generalizations, like the generalization that a person cannot be at two places at the same time, are also often left implicit. While this does decrease the complexity of the argument, it can be dangerous not to make the generalizations that justify one's reasoning steps explicit, as in such a case they cannot be checked for reliability and validity. This will be further discussed in Sections 3.1.3 and 5.2.1.

Multiple inferences can be chained to form more complex arguments with intermediate conclusions. Wigmore (1931, p. 13) called such chains of inferences "catenate inferences", Anderson and colleagues (2005) speak of "chains of reasoning" while Wyner and colleagues (2008) say such an argument can be seen as "a chain of reasons [or] reasons for reasons". In the example, we only have a testimony that someone *looking* like Haaknat got into the car. However, from this we can infer that Haaknat actually got into the car (see Fig. 3.3).

Anderson and colleagues identify the *interim probanda* in a chain of reasoning, the intermediate conclusions in a chain of reasoning from evidential data to the ultimate probandum. Here it must be noted that defeasible inferences (whether they

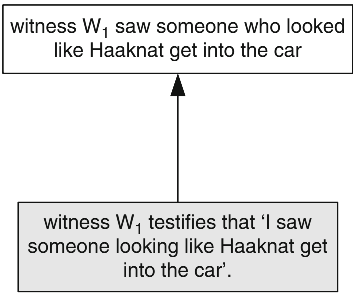



**Fig. 3.2** Argument with an implicit generalization

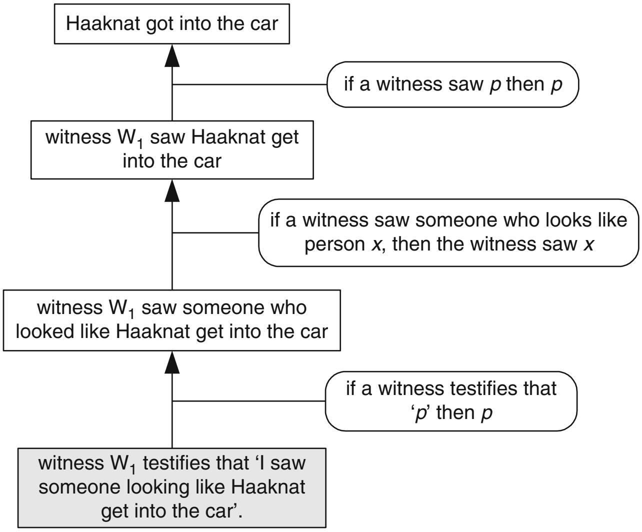



**Fig. 3.3** A chain of inferences

are combined with deductive inferences or not) are not transitive by definition. For example, consider the generalizations "students are usually adults" and "adults are usually employed in a full-time job". Knowing that Lucy is a student, we can fairly safely infer that she is an adult but we would not say that she is employed in a full-time job.

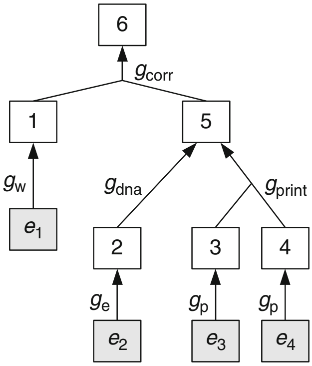
Arguments can also be represented as a linear sequence (Pollock, 1995, p. 87) where the order of the sequence represents the order in which the inferences were made. In the Haaknat example, when the police found the car near the park, they secured some hairs that were found on the seat of the car. Haaknat's DNA was also secured and analysed and it turned out there was a match between this DNA and the DNA profile of the hair found in the car. From this new evidence, we can infer that Haaknat had been in the car at some time.

(1) Expert testimony "the DNA profile of the hair in the car matches Haaknat's DNA" (*Input*)
(2) If an expert on DNA testifies that "DNA profiles A and B match" then DNA profiles A and B match (*Input*)
(3) The DNA profile of the hair in the car matches Haaknat's DNA (*1, 2*)
(4) If the DNA profiles of person *x* and a hair found at location *l* match, then *x* has been at location *l* (*Input*)
(5) Haaknat has been in the car at sometime (*3, 4*)

Behind each line of argument is noted whether the information stems from the "input" (the evidential data and accepted general knowledge, see Section 2.1) or whether the line is inferred from some other lines in the argument. Note that here an "expert testimony" generalization is used to infer a conclusion from the expert's testimony. So an expert testimony can be said to have its own associated generalization, just as we saw with a witness testimony in the argument in Fig. 3.1. Thus each type of evidential data (see Section 2.1.2) can be said to have its own associated generalization which is used when conclusions are drawn from the data. In this sense, Knigge (2001, Chapter 6) is right when arguing that form (the type of evidential datum, e.g. testimony) and content (the contents of the testimony) cannot be separated and that the type of evidential data and inference to the conclusion cannot be considered separately. The type of evidential data and the associated generalization are important when considering possible sources of doubt; for example, in Section 2.1.2 it was shown that with a witness testimony the veracity, objectivity and observational sensitivity have to be checked and that with expert testimony we could check if the expert's expertise is in the right field. The idea of recurring generalizations in evidential reasoning and their sources of doubt will be further discussed in Section 3.1.3.

Multiple arguments can be combined to infer a conclusion from more than one piece of evidence. In the Haaknat example, one could argue that the conclusion "Haaknat got into the car" is not particularly strong, as there may be other people who look like Haaknat. So if we want to infer that it was Haaknat who got into the car, we should ideally have more evidence for this. In addition to the hair found in the car, shoeprints made by shoes of the brand "Runner" were found in the car. When he was apprehended, Haaknat wore such shoes. So from this new evidence, we can infer that Haaknat had been in the car. This conclusion can be combined with the conclusion that someone who looked like Haaknat got into the car (Fig. 3.4 on p. 40). In this figure, multiple arguments are combined in a graph or tree like structure which shows many similarities to a Wigmore chart. For example, the evidential data is shown at the bottom and the conclusions at the top. Such a graph is also variously called an *argument graph*, *argument diagram* (Reed et al., 2004) or *inference graph* (Pollock, 1995). The graph can be seen as one argument for the conclusion that Haaknat was the man who got into the car. The argument consist of several *subarguments*; for example, the argument in Fig. 3.1 is a subargument of the total argument in Fig. 3.4. Similarly, the linear argument on p. 39 is also a subargument of the argument in Fig. 3.4. Bex, Prakken, Reed and Walton (2003) give a similar reconstruction of a Wigmore chart using linear arguments.

The argument in Fig. 3.4 shows two ways in which individual arguments can corroborate or accrue to strengthen a conclusion. In the case of $g_{\mathrm{corr}}$, there is a new generalization which arguably leads to a stronger conclusion than the previous one used to infer this conclusion, namely "if a witness saw someone who looks like person $x$, then the witness saw $x$". Prakken (2005a) calls this the "knowledge representation approach" to the accrual of arguments, where the accrual is "hand-coded" as a conditional with a conjunction of the accruing reasons in the antecedent. In the other instance of accrual in the argument in Fig. 3.4, two separate subarguments both lead to the same conclusion (5), namely that Haaknat was at some time in the car. Prakken calls this the "inference approach": after all relevant arguments based on individual reasons have been constructed, they are somehow aggregated.




$e_1$: Witness $W_1$ testifies that 'I saw someone looking like Haaknat get into the car'.
$e_2$: Expert $E_1$ testifies that 'the DNA profile of the hair in the car matches Haaknat's DNA'.
$e_3$: Police report: 'the shoeprints found in the car were of Runner shoes'.
$e_4$: Police report: 'Haaknat wore Runner shoes when he was apprehended'.

$g_w$: If a witness testifies that '*p*' then *p*
$g_e$: If an expert on DNA testifies that 'DNA profiles *a* and *b* match' then DNA profiles *a* and *b* match.
$g_p$: If a police report says that '*p*' then *p*
$g_{dna}$: If the DNA profiles of person *x* and a hair found at location *l* match, then *x* has been at location *l*
$g_{print}$: If the shoes of which the prints were found at location *l* are the same as person *x*'s shoes, then *x* has been at location *l*
$g_{corr}$: If a witness sees someone looking like person *x* get into a car and *x* has been in the car at sometime, then *x* got into the car

1: Witness $W_1$ saw someone who looked like Haaknat get into the car.
2: The DNA profile of the hair in the car matches Haaknat's DNA.
3: The shoeprints found in the car were of Runner shoes.
4: Haaknat wore Runner shoes when he was apprehended.
5: Haaknat has been in the car at some time.
6: Haaknat was the man who got into the car.

**Fig. 3.4** A complex argument graph for the conclusion that Haaknat got into the car

Exactly how these arguments should be aggregated will not be discussed here, as for current purposes it is sufficient to know that there are multiple ways of modelling corroboration or accrual.

In sum, the individual pieces of evidential data can be used to infer conclusions, the generalizations used can be made explicit and each step in the argument can be challenged (see Sections 3.1.2 and 3.1.3). The focus in this argumentative approach is on the individual items of evidential data, the conclusions that can be inferred from these items and the possible sources of doubt that arise when reasoning from a piece of evidence to a conclusion in this way. Because of this focus, the (Wigmorean) argument-based approach has also been characterized as the *atomistic approach* (see Twining, 2006, pp. 306–311; Malsch and Freckelton, 2009).

A second reason why reasoning with evidential arguments can be described as atomistic in nature is that their conclusion is most often not a complex set of propositions and their relations but rather a single proposition and thus multiple lines of reasoning based on various items of evidence in effect converge on a single conclusion. In most examples of Wigmore charts and evidential arguments, the conclusion is a single proposition like "Sacco was at the scene of the robbery and the shootings when they occurred" (Kadane and Schum, 1996, pp. 288–294) or "Jedrusik had a revengeful murderous emotion towards Umilian" (Wigmore, 1931, p. 64). This atomistic nature of arguments will be further discussed in Section 3.1.4.

### 3.1.2 *The Dialectical Nature of Argumentation: Attack and Defeat*

In the previous section it was shown how conclusions can be supported by evidential data trough arguments of varying complexity. Anderson and colleagues (2005) argue that in a complex argument, each step corresponds to a possible source of doubt. Such sources of doubt can be actively challenged by giving counterarguments that attack the original argument. The possibility of attack involves the *defeasibility* of the inferences: an argument for a certain conclusion can be overturned by new information which leads to, for example, an argument for the opposite conclusion or an exception to the application of a generalization.

In the literature (see Pollock, 1995), two types of attack are usually distinguished. An argument can be *rebutted* by giving a counterargument with as its conclusion the negation of a proposition in the original argument (Fig. 3.5). For example, say that Haaknat testifies he did not get into the car. From this, we can infer that Haaknat did not get into the car, the negation of the conclusion of the argument in Fig. 3.4. Note that rebuttal is a symmetrical attack relation: an argument with conclusion *p* attacks the argument with conclusion *not p* and vice versa.

Both Haaknat's testimony and the evidential data in the argument from Fig. 3.4 can be seen as *directly relevant evidence* for the question whether or not Haaknat got in the car (see Section 2.1.2). In other words, all the items of evidence that are leafs of the argument tree of which either *p* or *not p* is the root are *directly relevant* for determining whether or not *p* is justified. Note that it is also possible to rebut an intermediate conclusion or probanda. For example, an argument for the conclusion "the shoeprints found in the car were not of Runner shoes" attacks the intermediate conclusion 3 in the argument in Fig. 3.4. Prakken and Vreeswijk (2002) distinguish direct attacks and indirect attacks, where indirect attacks are directed against a conclusion of a subargument. Note that when one of the subarguments of an argument is attacked, the main argument is also attacked.

In addition to rebutting attacks, it is also possible to *undercut* an argument with another argument for why a particular inference is not allowed. That is, the proposition(s) in the argument are not negated but the inference step between propositions is. This is usually because the generalization that justifies the inference does not apply in the given circumstances because in the particular context there is an exception to the general rule. For example, say we have evidence that the expert who

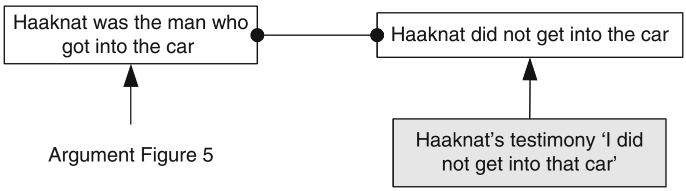



**Fig. 3.5** Two rebutting arguments

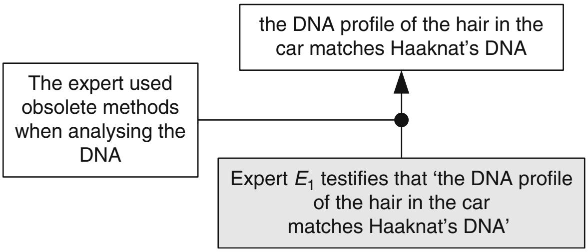



**Fig. 3.6** Undercutting an argument

analysed the DNA profiles used obsolete methods to analyse the DNA (Fig. 3.6, p. 42). Even though usually we would say that the opinion of experts on DNA can be trusted in this case we might argue that the opinion of an expert that uses obsolete methods cannot be trusted. Thus, the reason for believing that the DNA profile of the hair in the car matches Haaknat's profile is undercut but it is not denied that the two profiles match. Effectively, the defeasible inference from the source (in this case the witness testimony) is blocked. This attack relation is not symmetrical: the undercutting argument attacks the original expert testimony argument but not vice versa.

Evidential data that leads to an argument that undercuts an argument with as its conclusion *p* can be seen as weakening ancillary evidence for *p* or ancillary evidence against *p* (cf. Section 2.1.2). The second expert testimony that the first expert used obsolete methods is an example of such weakening ancillary evidence for the conclusion that the DNA profiles matched. Strengthening ancillary evidence can also be modelled using arguments (see Fig. 3.9).

In sum, there are multiple ways of attacking an argument (Fig. 3.7). Each attack effectively corresponds to what the Anderson and colleagues call "exposing a source of doubt in the argument". Merely attacking an argument, however, does not guarantee the argument's defeat. In order for an attacking argument to defeat another argument, the attacking argument has to be stronger than the other argument. With each specific argument some measure of its *strength* or *probative force* can be

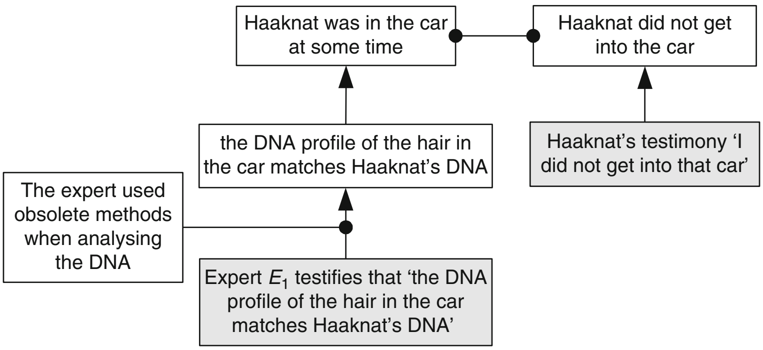



**Fig. 3.7** Attack relations in an argument graph

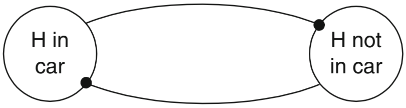
associated. The propositions as well as the inferences in the argument have their associated strengths, denoting the degree in which we are justified in believing such a proposition or inference. As was argued, the existence of evidential data cannot be denied so this data will have the highest available strength and consequently the data cannot be defeated. This strength can be expressed as a numerical probability (e.g. “the probability that datum *D* exists is 1”) but often it is also expressed in words (e.g. “it is certain that datum *D* exists”). Anderson et al. (2005, p. 230) give an overview of the different ways in which such probabilities can be expressed.

Calculating strengths of inferences and intermediate conclusions is not something that is easily done. The degree of support that a piece of evidence gives to some conclusion depends on the strength of the inference linking the evidence and the conclusion. In turn, the strength of this inference depends on the degree of belief we have in the generalization that justifies the inference. The first problem here is that this degree is often not easy to express in numbers. For example, if we say the witnesses under oath *usually* speak the truth, can we believe that they speak the truth 75% of the time? Or perhaps 85% of the time? Furthermore, should the degree of support be the product of the strength of the evidence and that of the inference or should we perform another operation on the two strengths? Such difficulties only increase when we consider complex arguments, where a chain of inferences leads to a conclusion or where two subarguments accrue.

Because of such difficulties, it is usually easiest to express the strength of arguments or generalizations relative to each other, that is, they express a preference for one generalization or argument over the other without mentioning the exact strength of either generalization or argument. Prakken and Sartor (1997) have modelled preference relations between generalizations (they call them “rules”). For example, we could argue that “arguments based on eyewitness testimonies by Henry are more reliable than arguments based on eyewitness testimonies by Bart, because Bart wears glasses and Henry does not”. In Prakken and Sartor’s framework, this would be modelled by saying that the witness testimony generalization for Henry (i.e. if Henry says “*p*” then *p*) is stronger than the witness testimony generalization for Bart and that hence an argument based on “Henry’s generalization” is stronger than an argument based on “Bart’s generalization”. Amgoud and Cayrol (2002) have also modelled preference relations but in their work the preferences are between arguments as-a-whole in an argumentation framework (see below) rather than between individual generalizations. Preference relations can themselves become the subject of discussion. For example, given new evidence that Henry is wearing contact lenses, the above preference relation between “Henry’s generalization” and “Bart’s generalization” can be attacked in Prakken and Sartor’s framework for defeasible priorities. Similarly, Modgil (2007) has defined a “meta-level”, on which it is possible to reason about the preferences between arguments.

After it has been determined which arguments are stronger than others and thus which arguments defeat which other arguments, the *dialectical status* of arguments can be assessed (see Prakken and Vreeswijk, 2002, Section 4). In this respect, arguments can be classified into three kinds: the *justified* arguments (those that survive

**Fig. 3.8** Mutually attacking arguments




the competition with their counterarguments), the *overruled* arguments (those that lose the competition with their counterarguments) and the *defensible* arguments (those that are involved in a tie). Say that, for example, we have two rebutting arguments that attack each other, viz. Figure 3.8.

Following Dung (1995) the above figure only shows the attack relations between two arguments and the internal structure of the arguments is abstracted from. This allows for a nice overview of the different arguments and their attack relations. Now, if no preference for any argument is defined, they are both defensible. If, however, for some reason, we prefer the left argument for "Haaknat was in the car", this argument is justified and the right argument is overruled.

The dialectical status of an argument depends on its interactions with all other available arguments. An important phenomenon here is *reinstatement* (Fig. 3.9): suppose that argument *B* defeats argument *A* but that *B* is itself defeated by a third argument *C*; in that case *C* reinstates *A*. In the example, say that we have an argument attacking "Haaknat in car" (e.g. the undercutting argument from Fig. 3.6). Because the leftmost argument is itself not attacked, it is justified and (assuming it is strong enough) defeats the argument for "Haaknat in car", which is in turn overruled. The right argument for "Haaknat not in car" is also justified because its only attacker is overruled. The notion of reinstating arguments allows for the modelling of strengthening ancillary evidence.For example, say that in the case of the arguments in Figs. 3.6 and 3.9 we have evidence that the expert is employed at a highly modern and respected institute where the newest methods and equipment are used. This evidence effectively "defends" the original argument for the conclusion that the DNA profiles match against any attacks by other arguments stating that the expert used obsolete methods or that the expert is some fiddler who has no knowledge of DNA-analysis techniques. At the same time, this new evidence is not directly relevant for the conclusion that the profiles match, so it can be seen as ancillary evidence that strengthens the expert's position.

Consequently, there are essentially two ways in which arguments can support or attack a conclusion. As was already argued above (Fig. 3.5), an argument *directly supports* a proposition *p* if there is a chain of inferences from the conclusion of the argument to *p*; for example, the left argument in Fig. 3.9 directly supports the

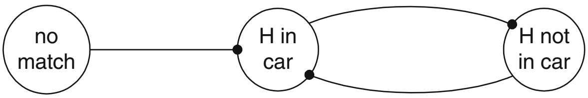



**Fig. 3.9** Reinstatement of arguments

conclusion of the argument in Fig. 3.4 because the first is a subargument of the second. Analogously, an argument *directly attacks* a proposition *p* if there is a chain of inferences from the conclusion of the argument to *not p*. An argument *indirectly supports* a proposition *p* if it attacks an attacker of another argument with the conclusion *p*. A set of arguments that defends itself against incoming attackers can be seen as a coherent and defendable position.[2] Finally, an argument *indirectly attacks*a proposition *p* if it undercuts an argument with the conclusion *p*.

Recall from the first paragraph of Section 2.3 that a proposition can ideally stand against criticism in a dialectical inquiry. In other words, an argument can be accepted if it cannot be successfully challenged in a *properly conducted dialogue*. In this respect, Toulmin (2003) presents his view of "logic as generalized jurisprudence": a logic for arguments should provide the essentials of a general rational process for analysing arguments just as jurisprudence provides the essentials of the legal process. The procedural and dialogical component of argumentation and reasoning in general has been presented more explicitly presented by, among others, Rescher (1977); Loui (1998). The structure and rules of a proper dialogue have been provided in the literature by, for example, van Eemeren and Grootendorst (2004). Their pragma-dialectic theory, which is meant to be used to analyze and evaluate argumentation in actual practice, provides the structure of a typical discussion and defines what can be argued at each stage of the discussion.A set of rules for a properly conducted discussion is also provided. For example, the *obligation-to-defend rule* states that "discussants who advance a standpoint may not refuse to defend this standpoint when requested to do so" and the *relevance rule* states that "standpoints may not be defended by non-argumentation or argumentation that is not relevant to the standpoint". Another example of work that provides rules for a properly conducted dialogue is (Walton and Krabbe, 1995). In formal argumentation, rules for conducting a proper dialogue have been incorporated into formal *protocols*. Many such protocols have been developed for persuasion dialogues (for an overview see Prakken, 2006), but protocols for negotiation dialogues are also given in the literature. Such informal and formal protocols serve not only an analytic function (is the discussion rational in that it follows the rules?) but also a heuristic function (what are our options if we want to conduct a rational discussion?).

The basic idea of using a dialogue as a means of rationally analysing an argument is best explained by taking an example from simpler and more mathematically inclined dialogues. A number of authors (e.g. Vreeswijk, 1993; Loui, 1998; Prakken, 1997) have explicitly modelled the analysis of arguments in a procedural way.[3] The basic idea of these so-called *dialogue games* is that a simple game is played between two players, a proponent and an opponent. The proponent starts by moving an argument that needs to be tested and each subsequent move (by either

---

[2]In Chapter 5, the idea of a set of arguments "defending" itself against attackers is made more clear through a series of formal definitions.

[3]In Chapter 5, a formal dialogue game for a hybrid argumentative story-based theory will be provided.

the proponent or the opponent) contains an argument that attacks an argument of the other player. The rules of the game determine, for example, whether a player may repeat his earlier moves or whether a player may move only undercutters or rebuttals. Say, for example, that we have a game in which the proponent starts by moving an argument and may not repeat his moves, all of the arguments in the opponent's move must defeat the proponents move and the arguments in proponent's subsequent moves must undercut the opponent's arguments in the previous move.[4] The proponent now starts by moving the argument "Haaknat not in car". The opponent must attack this argument and hence moves "Haaknat in car". The proponent must undercut this argument; this can be done by moving "no match". There are now no more arguments and the opponent has no more valid moves so the proponent wins. Here, it can be seen that the argument game essentially provides a dialectical proof theory for arguments: the initial argument can be said to be (defeasibly) provable if the opponent can attack (and defeat) each move the opponent makes. In the current game, "Haaknat in car" is then not provable: the opponent can attack this argument by moving "Haaknat not in car" and, seeing as the proponent may not repeat his moves, he cannot attack this argument so the opponent wins.

### 3.1.3 Generalizations and General Knowledge in Evidential Reasoning

Generalizations from our stock of shared knowledge play a pivotal role in reasoning with evidential arguments. They can warrant inferences from the evidential data to conclusions and thus can be seen as the glue that keeps an argument together. Generalizations are not always of the conditional, "warranting" type: other general assumptions of a non-conditional form are also generalizations. In Section 2.1.3 on general knowledge, the dangers of generalizations were briefly discussed. These dangers of generalizations can be lessened by specifying exactly which generalizations we use, how we use these generalizations and from which sources the generalizations stem. Making the knowledge from our shared stock explicit in this way minimizes the chance that it is misunderstood or used in the wrong way. The more explicit the knowledge, the more it is open to analysis and criticism. In this way, an actual cognitive consensus about the generalizations we use can be reached through a critical dialectical process as specified at the end of Section 3.1.2.

Cohen (1977) assumes that there is no discussion about the kinds of generalizations used in legal reasoning and that the only discussion is on whether or not a certain generalization is applicable in a particular case. While Anderson and colleagues make a valid point that such a consensus cannot be readily assumed in a diverse society, I would argue that there are quite a few generalizations that, in one way or another, are consistently used by all kinds of reasoners in all stages of the

[4] This argument game is essentially an adapted version of (Prakken, 1997).

process of proof. When looking at evidential reasoning (or indeed at reasoning in general), one sees that many arguments, as well as attacks on them, are instances of recurring patterns, such as inferences from witness or expert testimonies.

Bex and colleagues (2003) have argued that in this sense, *argumentation schemes* play an important role in reasoning with evidence. Argumentation schemes are forms of argument that represent stereotypical patterns of human reasoning in a conditional form, just like generalizations. The idea of defining recurring patterns of reasoning through argumentation schemes or generalizations is the subject of much current study in argumentation theory, artificial intelligence and law. Walton has applied reasoning with argumentation schemes to a wide variety of different subjects; some examples of these subjects are general reasoning with evidence (Walton, 2002) and reasoning with and about motives in criminal cases (Walton and Schafer, 2006). Other authors (e.g. Bex et al., 2003; Verheij, 2003b; Gordon et al., 2007) have also made extensive use of argumenttation schemes in more formal, AI-oriented work. For a general overview of reasoning patterns, the reader is referred to (Walton et al., 2008, which is essentially an update of Walton, 1996). As an example of an argumentation scheme, take the well-known scheme for argument from expert opinion (Walton et al., 2008, p. 244):

> Source *e* is an expert in domain *d*.
> *e* asserts that proposition *a* is known to be true (false).
> *a* is within *d*.
> *Therefore*, *a* may plausibly be taken to be true (false).

The notion of argumentation schemes is obviously very closely related to the notion of generalizations: argumentation schemes are conditional rules based on world knowledge which can be used to draw inferences. The above argumentation scheme is a slightly more general version of the DNA expert generalization given on p. 39. In the generalization, the domain *d* is “DNA analysis” and it is assumed that “profiles *a* and *b* match” is a statement within the domain of DNA analysis.

Recall that in Section 3.1.1 it was argued that each type of evidence has its own associated generalization which allows us to draw inferences from that particular type of evidence. In this way, the various types of evidence point to generalizations that are often used in reasoning with evidence. Above we already saw the generalizations for witness testimony and expert testimony. Another example is the generalization for inference from documentary evidence: “if a document states that ‘*p*’ then (presumably) *p*”. In this way, we might accept stereotypical ways of reasoning about which there is a consensus, at least in the legal and philosophical community, and thus accept that there are certain valid generalizations that can be used in rational reasoning about evidence. Because each type of evidence has its own associated generalization, the law may also point us to generalizations which are accepted by default; for example, in Dutch law witness testimonies are explicitly stated as a species of evidence on the grounds of which a judge can form his decision (article 339 paragraph 3 and article 342 DCCP). This means that it is highly unlikely that the legislator believed the witness testimony generalization to be false by default. Other sources of generally accepted patterns of reasoning are the argumentation schemes as proposed by the above-mentioned authors and Pollock (1995), who presents a number of stereotypical patterns of epistemic reasoning; for example, rules for reasoning from perception or memory.

One of the main points of looking for stereotypical patterns of reasoning is that for each generalization, some typical sources of doubt can be given. For example, the witness testimony generalization can be undercut with arguments questioning the witness' veracity, objectivity and observational sensitivity. Pollock (1987) defines a standard undercutter for his perception rule as follows: "the present circumstances are such that having a percept with content *p* is not a reliable indicator of *p*" and this undercutter can be used to question a witness' observational sensitivity. In the same way, the expert testimony generalization has several typical sources of doubt. Walton provides each argumentation scheme with a number of *critical questions*, which point to possible sources of doubt in an argument based on the scheme. These critical questions also fit into the dialectical view on argumentation as they can be used in a question-and-answer dialogue. The following six basic critical questions matching the appeal to expert opinion have been recommended in (Walton et al., 2008, p. 246):

(1) *Expertise Question*: How credible is *e* as an expert source?
(2) *Field Question*: Is *e* an expert in *d*?
(3) *Opinion Question*: What did *e* assert that implies *a*?
(4) *Trustworthiness Question*: Is *e* personally reliable as a source?
(5) *Consistency Question*: Is *a* consistent with what other experts assert?
(6) *Backup Evidence Question*: Is *a*'s assertion based on evidence?

Answers to these critical questions can lead to various types of counterarguments. For example, a negative answer to the "field question" would undercut an argument from expert opinion and a negative answer to the "consistency question" points to a possible rebutting counterargument with an opposite conclusion. The undercutting attack in the example in Fig. 3.6 (p. 42) is an answer to perhaps another critical question of the form "Did the expert use the right method to determine the truth of *A*"?

One argumentation scheme that plays an important role in reasoning in the process of proof, particularly in the context of discovery, but which is not related to a specific piece of evidence is the *argument from sign*. Walton's (1996) example is as follows: "Here are some bear tracks in the snow, therefore a bear passed this way". The generalization based on this argumentation scheme can be phrased as "If *P* is a sign for *Q* and *P* is observed, then *Q* can be assumed". Eco(in Eco and Sebeok, 1983) argues that there are various types of signs. They can be things such as shoeprints, fingerprints and hairs which have a more or less direct correspondence with their cause (i.e. a shoe or a finger) but also more general clues such as the observation that "someone is of a medical type but has the air of a military man".[5] A general critical question for the argumentation scheme from sign is "Could the sign have been caused by *P*?"

[5] Sherlock Holmes used these signs to explain that Watson was an army doctor (In Conan Doyle's *A Study in Scarlet*).

An interesting argumentation scheme in reasoning about motives and actions in the process of proof is the *abductive practical reasoning scheme* as proposed by Bex et al. (2009). This abductive scheme[6] makes it possible to infer an explanation for a current state of affairs in terms of a motivated action:

> The current state of affairs is *C*.
> *G* is a goal for agent *a* motivated by motive *M*.
> Doing action *A* in some past state *P* would bring *a* in the current state *C* in which *G* is reached and thus *M* is fulfilled.
> *Therefore* agent *a* has probably performed action *A* in state *P*

This scheme is based on Atkinson and colleagues' (2006) extension of Walton's original scheme for normal, non-abductive practical reasoning, which enables argumentation about what action should be taken in the future by some agent to further his values and goals. In the abductive scheme, motives are similar to these values, abstract principles which an agent may or may not want to promote (see Section 2.3.3 on motives). Consider the Haaknat example; the current state of affairs is that the supermarket has been robbed. Say that Haaknat had as an underlying motive that he wanted to satisfy his addiction and that, because he needed money to fulfil this motive, he had as his goal to rob the supermarket. Now, Haaknat could reach this goal by actually robbing the supermarket and (provided that he got some money out of the robbery) thus fulfil his basic motive. Therefore, Haaknat has probably robbed the supermarket. Notice that this scheme assumes that Haaknat had a certain motive and hence certain goals.Critical questions for this scheme allow us to question these assumptions, for example, "is the motive legitimate?" or "is there another possible motive for robbing the supermarket?". Other critical questions concern the physical explanation itself, for example, "Given his motives, could Haaknat have reached the current state by doing a different action?". The abductive practical reasoning scheme and its associated critical questions can thus be used to explain why and with what motives an agent took certain actions in the past and to critically analyse the explanation.

Argumentation schemes and their critical questions point to stereotypical patterns of reasoning and ways to criticize these patterns. In Section 5.2.1, the role of argumentation schemes in defining the contextual logic for evidential reasoning will be briefly discussed and there some more examples of typical inferences and their sources of doubt will be given.

Ideally, a generalization comes from a clearly defined source, so that if we are doubtful about whether the generalization should be believed we can check the original source. This source can be an expert or a (legal) document. For example, if we learn from a medical book that appendicitis often leads to pains in the side, we might believe the generalization that heavy pains are evidence for appendicitis. Often however, generalizations will not be phrased as a neat conditional rule. For example, consider the reasoning in Fig. 3.10, where it is argued that witness rule follows from the Dutch Code of Criminal Proceedings.

---

[6] In order to have a similar-looking scheme as the one for expert opinion given above, the abductive practical scheme presented here is of a slightly different form than the one in (Bex et al., 2009).

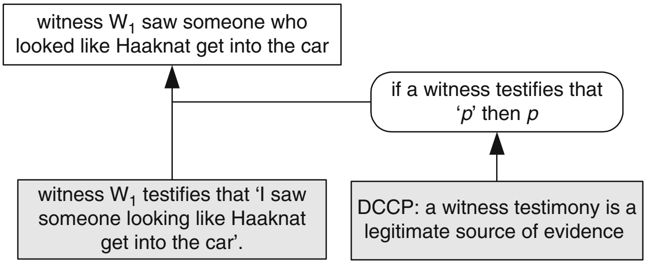



**Fig. 3.10** Evidence for a warranting generalization

Note the similarity between the above structure and Toulmin's (2003) scheme for the layout of an argument. Here, the witness testimony is the datum, which is the basis of the claim that the witness saw someone who looked like Haaknat. The generalization acts as the warrant and the DCCP as the backing, showing why the warrant holds.

Generalizations, however, often do not stem from a clearly defined source but rather from our stock of knowledge; in such a case, they are usually backed by experience or general knowledge. In previous work (Bex et al., 2003), we argued that experience-based generalizations seem to be based on a commonsense counterpart of scientific induction and that reasoning from a "general knowledge source" can be formulated as a new generalization: "It is general knowledge that '*p*' is evidence for *p*". Possible undercutters of this generalization are that a piece of general knowledge is infected by prejudice or value judgement. In Fig. 3.11 this reasoning is visualized.

Notice that in Fig. 3.11 the proposition from which the generalization is inferred is rendered as a white box, which means that it is not considered to be evidential data (the existence of which cannot be challenged) but rather a general knowledge assumption which can itself be called into question (i.e. attacked by an argument). In a sense, this general knowledge assumption is in itself a generalization about what is contained in our stock of knowledge. Thus, non-conditional generalizations

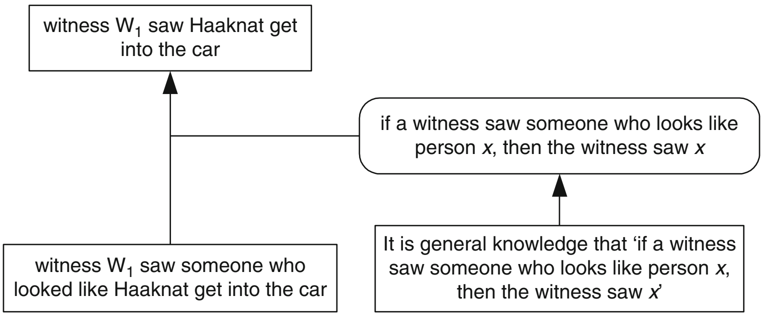



**Fig. 3.11** Generalizations as general knowledge

can be used to support (i.e. act as the backing for) other generalizations.If we model reasoning with generalizations as in Fig. 3.11, there are essentially four ways to attack a generalization (adapted from Bex et al., 2003):

(1) *Attacking the validity of the source of the generalization*: it is not general knowledge that "If a witness saw someone who looks like person *x*, then the witness saw *x*".
(2) *Attacking the defeasible derivation from the source*: it is indeed general knowledge that if a witness saw someone who looks like person *x*, then the witness saw *x*, but this particular piece of general knowledge is based on a belief from folk psychology that people are always accurate at recognizing faces.
(3) *Attacking application of the generalization in the given circumstances*: Usually it is true that "If a witness saw someone who looks like person *x*, then the witness saw *x*". However, in this case we cannot conclude that the witness saw Haaknat as Haaknat has a very common appearance.
(4) *Attacking the generalization itself*: it is not the case that "If a witness saw someone who looks like person *x*, then the witness saw person *x*".

Note that the first type of attack is only possible if the source of the generalization is not evidential data but rather an assumption from the stock of knowledge. For example, it is not possible to deny the source from Fig. 3.10 by arguing that "a witness testimony is not a legitimate source of evidence". In this case it is possible to attack the derivation from the source. The main difference between attacks of the third and the fourth kind is that the third kind of attack accepts the generalization as a general rule but denies its application in the case at hand, while the fourth kind of attack denies the generalization as a general rule ("it is not the case that usually…"). Figure 3.12 sketches how these various attacks can be modelled as attacks on an argument. Here it can be seen that various ways of arguing about and with generalizations can be modelled in the argument-based approach.

Another issue concerning (conditional) generalizations in the argument-based approach is the *refinement* of generalizations. Following earlier work by Loui and Norman (1995), Bex and Prakken (2004) have shown that this refinement can

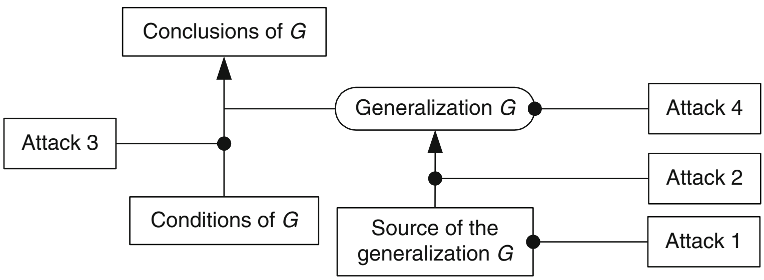



**Fig. 3.12** Attacking a generalization

be done in essentially two ways. First, a generalization can be "unpacked". With unpacking, a single-step argument based on a particular generalization is replaced by a chain of reasons based on multiple generalizations with the same start and end point as the original argument. Kadane and Schum (1996) give an example of unpacking the witness testimony generalization. Because this generalization has the features of veracity, objectivity and observational sensitivity, it can be rephrased as three separate generalizations. The original generalization "if a witness testifies that '*P*' then usually *P*" could thus be unpacked into "if a witness testifies that he observed *P* then usually he believes that he observed *P*", "if a witness believes that he observed *P* then usually his senses gave evidence of *P*" and "if a witness" senses gave evidence of *P* then usually *P*'. An argument based on the original generalization can thus be unpacked, viz. Figure 3.13.

The second type of refinement involves exposing hidden conditions of generalizations. For example, "if two witnesses testify that '*p*' then *p*" can be argued to have an additional condition that the witnesses did not confer. The rationale here is that witnesses who confer can possibly change their testimonies so that they are compatible. We can make this explicit by changing the generalization into "if two witnesses testify that '*p*' and the witnesses did not confer then *p*". The original argument is then changed as in Fig. 3.14 . Notice that the new condition does not necessarily follow from evidential data. Ideally this will of course be the case, so here we should look for further evidence that the witnesses did not confer. An interesting observation is that a rebuttal of such a hidden condition can in a way be seen as a possible undercutter of the original argument: arguing that the witnesses did

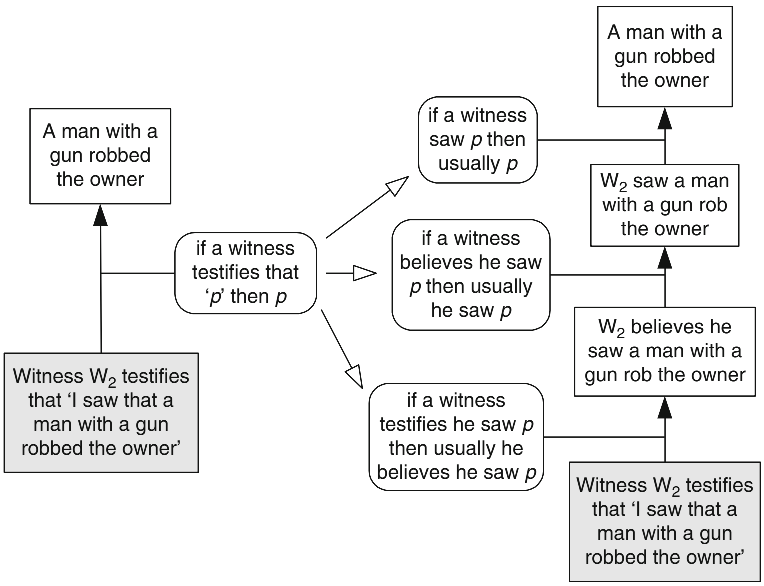



**Fig. 3.13** Unpacking an evidential generalization

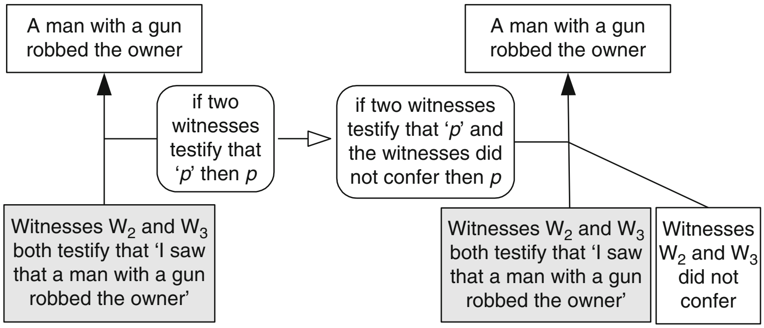



**Fig. 3.14** Adding conditions to an argument

somehow confer will undercut the original argument but rebut the assumption of the refined argument. Hence, by exposing hidden conditions we make possible exceptions to the generalization explicit. This second way of refining generalizations is close to what Crombag et al. (1994) mean with "making a generalization safe". They say that in the anchoring process, generalizations are made increasingly detailed until the reasoner is satisfied that most important exceptions to the generalization have been made explicit. Refining an argument in this way and then supporting the new condition with evidential data seems to be very much like giving strengthening ancillary evidence; after all, if we have evidence that the witnesses did not confer, the inference to the conclusion is stronger than if we just assume they did not confer by default.

### *3.1.4 Summary and Evaluation*

This chapter has introduced arguments as a general form of reasoning. It was shown how complex arguments or argument graphs can be built and how these arguments can be attacked and defended against attackers. It was shown how the probanda can be inferred from the evidence in a case and how generalizations can be analysed in various ways. This section evaluates the use of arguments and argument graphs in reasoning with criminal evidence. First, some specific uses of arguments in the three contexts of the process of proof (discovery, pursuit and justification) will be briefly discussed. The atomistic and evidential nature of arguments has certain advantages and disadvantages and these will also be discussed below. This section ends with briefly discussing the naturalness and rational well-foundedness of the argument-based approach.

In the context of discovery, evidential arguments can be used to generate simple hypotheses. The argument from sign can be used to infer causes for single clues of tangible evidence such as hairs, blood, bullet casings and so on. The abductive practical argumentation scheme allows us to derive simple explanations for a state of affairs consisting of single motivated actions. Other more general abductive argumentation schemes (see, for example, the general abductive scheme in Section 2.3.1) can also be used in fact finding, as Schum (2001) has persuasively argued.

In the context of pursuit, evidential arguments allow us to prove the individual probanda in a case by supporting them with evidential data. In this way, the relevance of evidential data for a particular conclusion can be shown. The possibility to attack probanda with counterarguments allows them to be tested in a rational dialectical process. Furthermore, undercutting an argument by arguing that a generalization is not applicable in the case at hand or arguing against the validity of a generalization in general allows for the critical analysis of the chain of reasoning from the evidence to the probanda. Attack graphs such as the one shown in Fig. 3.7 can be used to determine the total force of evidence. As the example shows, the status of one argument may depend on attack relations any number of other arguments. Thus, the attack graph might show that, for example, the testimony of one witness is important because it reinstates a large number of other arguments for a particular position. Such attack graphs may also be useful in the pursuit of further evidence, as they give an overview of which arguments need to be defeated if we want the arguments supporting our preferred hypothesis to be justified.

Evidential arguments can be used in the process of justification by showing the exact chain of reasoning from evidence to probanda. Thus argument graphs help not only the reasoner but also third parties to understand the reasoning and identify sources of doubt. A graph showing the attack relations between arguments based on evidence can aid by giving an overview of how the various pieces of evidence interact and it shows why a particular combination of arguments should be believed.

Recall that reasoning with arguments has been called *atomistic* because the various elements of a case (i.e. probanda, evidential data) are considered separately and the case is not considered "as a whole". This atomistic nature of arguments ensures that the various pieces of evidence are clearly individuated and that the evidential generalizations that justify the reasoning steps from the data to the probanda are explicitly mentioned in the argument. Furthermore, sources of doubt in a chain of reasoning can be made explicit by refining arguments in various ways. Generalizations and their sources can be tested in multiple ways and the dialectical process of attacking and defending arguments allows for a proper discussion not only of the evidential data but also of the general commonsense knowledge so that a proper cognitive consensus about the stock of knowledge can be reached. In sum, the atomistic nature of arguments makes them very useful for carefully analysing each piece of evidence, the probanda and the general knowledge used in reasoning from this evidence to the probanda.

However, due to this atomistic nature of arguments the overview of the case tends to be lost in a purely argument-based approach. In a case, the various hypotheses about what (might have) happened are usually not simple propositions but rather hypothetical scenarios detailing the motives and actions of each of the persons involved in the case. Such a scenario is a story, a coherent set of events (see Section 3.2). The conclusion of an evidential argument is usually a single element of such a scenario, an individual state or event (e.g. "Haaknat got into the car") and the hypothetical scenarios about "what happened" in the case are thus cut into pieces. Take as an example the Haaknat case. The question here is "why is Haaknat the robber"? This ultimate probandum can be supported by providing the arguments for each of the penultimate probanda which together roughly form our hypothesis about what happened. These penultimate probanda that form the main elements of this hypothesis are that Haaknat had a clear motivation for robbing the supermarket, that he could have been the robber (because he was seen near the supermarket just after it was robbed) and that he acted suspiciously when he was found in the park.

In Fig. 3.15 there are three separate subarguments each supporting one element of the hypothesis that Haaknat was the man who robbed the supermarket. Now, to combine these three elements we need a generalization of the form "if a supermarket is robbed and there is a man who has a clear motive for robbing a supermarket and this man could have been the person robbing the supermarket (based on his whereabouts just after the robbery) and this man was acting suspiciously when apprehended then this man is the person who robbed the supermarket". This generalization is fairly contrived; furthermore, it does not capture the fact that reasoning with general knowledge is often not done using individual statements but rather using more holistic schemes (see Section 3.2.2 on story schemes).

The atomistic nature of arguments also has its effect on the usefulness of arguments in all three contexts of the process of proof. In the context of discovery, relatively simple signs and clues can be interpreted using, for example, the argument from sign and hypotheses consisting of a motive and an action can be constructed using abductive practical arguments. However, an argument from sign only allows for the inference of a single cause of a clue and an explanation inferred with the abductive practical scheme is about just one action and its motives; using these argumentation schemes it is not possible to infer complex hypotheses consisting of multiple related events that portray the incident and the circumstances under which it took place. Similarly, in the context of pursuit one cannot compare

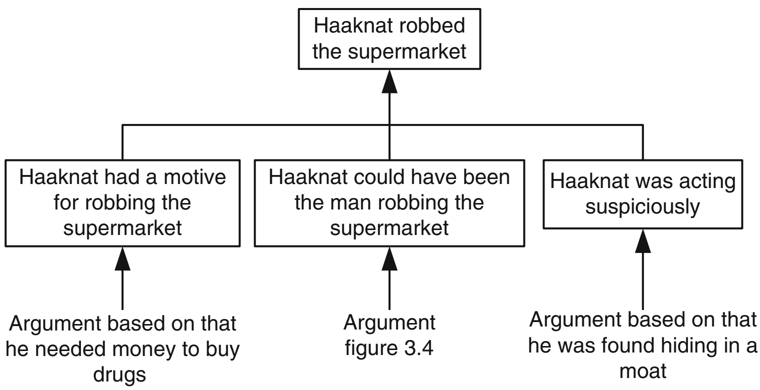



**Fig. 3.15** The penultimate and ultimate probanda in the Haaknat case

different accounts of "what happened", because such accounts are not provided in the argumentative approach. Furthermore,with respect to an explanatory justification, presenting a third party with a set of complex arguments for and against particular probanda does not help them understand exactly what happened and under what circumstances the crime happened. When presented with such arguments for the separate elements of a scenario, we are in a sense only given the "pieces of the puzzle" and without an idea of what the eventual image should look like, it can be very hard to put together these pieces. In the work on the Anjumer murder case (Crombag and Israëls, 2008, see the case study in Chapter 6), it is noted that in the judgement by the Court of Appeals only the evidential data that the Court used to come to its conclusion is mentioned. In the judgement, no general scenario of what happened in the case is given and it is left up to the reader of the judgement to cut-and-paste the various conclusions that may follow from the evidence into a coherent story.

Reasoning in the argument-based approach can be characterized as *evidential reasoning* as all the example arguments and argumentation schemes presented in this section are all of the form "*e* is evidence for *c*". While some authors (e.g. Shanahan, 1989) have argued that reasoning with evidential generalizations can be counter-intuitive, others (e.g. Pearl, 1988a) argue that we cannot expect people to always express their ideas as a causal generalization. He argues that people also use evidential rules or knowledge about how familiar situations can lead to previously successful guesses, for example, that symptoms suggest conditions: whilst every doctor knows that appendicitis can cause pain in the side, depending on the situation some doctors would also say that "pain in the side" (symptom) is evidence for "appendicitis" (condition). Especially when reasoning directly from a source of evidence, it seems that it is most natural to reason from the evidential data to a conclusion. In the judgements of the Dutch criminal cases I have studied, reasoning from evidence to some event is often of an evidential form. For example, in the judgements of the Nadia van der V. case (which was analysed in Bex and Verheij, 2009), the judges use phrases of the form "the event can be inferred from evidence $e_1$", "this event is based on (or supported by) evidence $e_1$". Research in our project (van den Braak et al., 2008) has shown that in the case of "testimonial knowledge", that is, information from testimonies and evidential documents, people find it significantly harder to interpret causal relations like "*x* bought a weapon *causes* witness *w* to testify that *I saw that x went into a store and came out with a weapon*" than they find it to interpret evidential relations like "Witness *w* testified that *I saw that x went into a store and came out with a weapon is evidence for x* bought a weapon". So it seems that reasoning with evidential generalizations as described here is quite natural, at least when we are reasoning with generalizations that allow us to infer conclusion directly from the evidential data, like the witness testimony generalization.

However, evidential generalizations may not always be the most natural way to express certain knowledge. Pearl's argument that people sometimes want to express their knowledge in an evidential way can also be said to hold the other way around, that is, people may also want to express some information in a causal way (also see

Simon, 2001). For example, in criminal cases issues like the cause of death are often expressed with causal generalizations. Furthermore, the exact causal structure of a hypothetical scenario and the (causal) relations between the various the elements of the scenario (e.g. the motives and actions) are also not shown and cannot be reasoned about.[7] Finally, when reasoning with just evidential generalizations it is not possible to perform predictive reasoning so possible consequences of accepting a specific hypothesis cannot be inferred. In the Haaknat example, it is for instance not possible to predict that Haaknat may have caused fingerprints to be on the steering wheel of the car and that these fingerprints should therefore be found. Thus the search for new possible evidence is also limited in the argument-based approach.

Argumentative reasoning provides us with a rationally sound way of analysing and assessing reasoning with evidence. The defeasible inference of conclusions from evidence and the possibility of constructing arguments and counterarguments allow for a rational process which, if conducted properly, provides a rationally justified conclusion.The main ideas on argumentation presented in this chapter are almost all logically and conceptually well-developed in the literature and a tradition of research on informal and formal argumentation provides the argument-based approach with the necessary academic grounding. Using data to infer a simple conclusion through evidential inference is a natural way of showing the support the data gives to the conclusion and argument charts allow for a careful and detailed analysis of the evidence. Research has shown that when people are forced to diagram their arguments, this improves their critical thinking abilities (Rider and Thomason, 2008; Twardy, 2004). However, van den Braak and others have published an overview of empirical research into argument diagramming (van den Braak et al., 2006) and many experiments were not valid or could not be tested for validity and even though the results often pointed in the right direction (i.e. that the reasoning improves through diagramming), the results were often not statistically significant. Furthermore, the argument-based approach seems less natural when it comes to *organizing* a mass of evidence. Properly organizing evidence so as to provide an overview is just as important in the process of proof as the detailed analysis of individual evidence is.

## 3.2 Stories

Stories are a popular form of communication between people. Books, newspapers, the 8 o'clock news, movies, a friend recounting his vacation, they all tell stories. Good stories can play a major role in our understanding of the world, as they structure information in a way that is easy to process for humans. Our memories, for example, consist (at least to the mind's eye) of stories or fragments of stories

---

[7] In the Umilian chart (Wigmore, 1931, p. 56, see Fig. 7.4 in this book), node 9 is effectively a cause for node 8: J falsely charging U with bigamy *caused* a murderous emotion (in U) towards J. However, the scenario is here not properly separated from the argument and this combination of evidential and causal reasoning should be treated with care (cf. Pearl, 1988a).

(Schank, 1986). A police detective, when trying to figure out "who did it" constructs a story around the available evidence and thus tries to find out what happened. People tell each other stories, for fun or when explaining something. Stories come in many different forms; some stories simply recount a series of events without giving any "colour" to the proceedings, while others explicitly or implicitly express certain opinions.

Aristotle was one of the first to discuss stories. He identified six elements which are vital to a traditional Greek tragedy (see Aristotle, 2005), examples of which are the Plot (the structure of events), Character (the characters and their goals and intentions in the story) and Spectacle (the costumes, stage etcetera). After Aristotle, linguistics and literary theory were the first disciplines to analyse stories and their elements. Most literary theorists consider *events* to be the most important building blocks of a story and the expressions "basic story structure" and "event structure" are used synonymously by many theorists. The Russian formalist Vladimir Propp (1968) tried to find the main types of events that compose a story. He analysed a corpus of 115 Russian fairytales and found that stories always start with an initial situation and that events (or *functions* as he calls them) are always one of 31 standard events and always appear in the same sequence. Furthermore, these sequences of events always lead to a climax or ending. For example, "The villain kidnaps the princess", "The hero and villain join in direct combat" and "The villain is defeated" always appear in this sequence and together compose a general beginning – action – ending structure of a story.

Following up on Propp's work, researchers from the cognitive psychology community (Johnson and Mandler, 1980; Rumelhart, 1975), who were mainly interested in story understanding and people's use of commonsense knowledge, developed a number of *story grammars*.These grammars were attempts at formalizing the structure and different parts of a typical story. Most of these grammars divide a story into episodes, each of which have a basic "setting – goal – action – consequence" structure. The idea behind the story grammars is that stories have a syntactical structure just as an individual sentence has a syntactical structure (object, subject, main verb etc.) and that information which is organized is easier to store in and recall from memory.

In later research the attention shifted towards story understanding by using a set of "general action sequences" orscripts (Schank, 1986; Schank and Abelson, 1977). While Schank and Abelson also use a basic episode scheme, they also argue that story understanding is more content-driven, in that more specific and detailed information about standard patterns of actions are used when reading and understanding stories. These standard patterns or sequences are modelled as scripts which we have stored in memory. The much quoted "restaurant-script", for example, contains information about the standard sequence of events that take place when somebody goes to dine in a restaurant. In order for a story about a restaurant to be understandable it does not have to mention all the events (e.g. ordering, eating, paying etcetera) because we refer to our memorized restaurant script to complete the story.

Since the early nineteen eighties stories also play an important part in theories on how judges, jurors or police investigators reason with the evidence in criminal cases.

Authors such as Bennett and Feldman (1981) and Pennington and Hastie (1986, 1988, 1992, 1993a, b)[8] argue that decision making in criminal cases is done by constructing stories about "what happened" using the evidence in the case and then comparing these stories, thus trying to find the best story. A good story should not only be compatible with the evidential data but it should also be well-structured and correctly describe a general pattern of states and events one expects to come across in the world. Crombag et al. (1994; Wagenaar et al., 1993) have a similar approach. However, they also found that in many cases a good or plausible story which is insufficiently compatible with the evidence wins over a bad or implausible story which is compatible with the evidence. To overcome this problem, Crombag and colleagues proposed their anchored narratives theory (ANT), according to which a story should be sufficiently anchored in reality using generalizations.

According to De Poot and colleagues (2004), stories and anchoring also play an important part in the investigative phase of a case, where stories serve as guidelines in the search for new evidence. Anderson et al. (2005; Schum, 2005; Twining, 1999, 2006) also say that stories can play a role in the context of discovery. Furthermore, like Wigmore (1931, pp. 659–660), they also maintain that stories are psychologically necessary in the determination of the facts of a case, in that a story is used to organize and present the evidence in such a way that it is easily understandable. Like Crombag and colleagues, they point to the dangers of stories and develop a protocol for analysing the plausibility and evidential support of stories.

As opposed to the argument-based approach to reasoning with evidence, which is mainly based on a combination of Wigmore charts and ideas from informal and formal argumentation theory, it is hard to speak of one distinct approach to reasoning with stories and the story-based approach has been less developed over the years. The story-based approach as discussed here is mainly based on ideas by Pennington and Hastie and Crombag, van Koppen and Wagenaar combined with formal work on causal-abductive model-based reasoning such as Thagard (2004) and Josephson (2002) and work in the legal field by Pardo and Allen (2007).

### *3.2.1 The Causal Structure of Stories*

A story is essentially a particular, coherent and chronologically ordered sequence of states and events (Bal, 1985; Toolan, 2001; Anderson et al., 2005). It is not always possible to maintain a strict chronological ordering; for example, when two events happen at the same time, it makes sense to first describe one event and then the second. Here the distinction is between the *story* and the *discourse* should be noted (Toolan, 2001). The story is the sequence of events and the discourse is the way in which a story is presented and through what medium (i.e. images, text, figures, film). Changing the way a story is presented does not have to change the story itself. For example, telling a story through flashbacks does not change the basic

[8] In the remained of this chapter Pennington and Hastie will often be referred to without explicitly mentioning one of these articles.

(chronological) event structure. In the research on stories, the combination between story and discourse is often called a *narrative*.

Not just any chronologically ordered sequence of events is a story: in order to be a story, a sequence of events should also be *coherent* in some way. First, the story should not contain obvious contradictions. For example, a story where a person was at different places at the same time is obviously not coherent. Second, the story should adhere to an (implicit) causal structure. Take, for example, the following sequence of events:

> Haaknat is a drug addict – the supermarket is robbed by a masked man – Haaknat jumps into the moat

Even though this is a chronological sequence of events, we would not say that it is a coherent story. This is because the events in the above sequence are seemingly not *causally connected*. Research in cognitive psychology (e.g. Trabasso and Sperry, 1985; Trabasso and van den Broek 1985) has shown that the states and events in a coherent story should all be connected through some sort of (implicit) causal chain. Further research in story grammars and story understanding (e.g. Schank and Abelson, 1977; Mandler and Johnson, 1977) assumes that in the underlying structure of stories the events are connected by a combination of causal and temporal relations.

However, the exact causal relations between the events in a story are often left implicit when a story is told. Take, for example, the following story about Haaknat robbing the supermarket:

> John Haaknat is a drug addict who is desperately in need of money, so he decides to rob the local supermarket. Haaknat parks his red car around the corner from the supermarket and puts on a balaclava. Haaknat then goes into the supermarket and threatens the owner at gunpoint. The owner hands him the money and Haaknat walks back to his car, taking off his balaclava when he is just around the corner. He gets into his car and takes off. Just as he drives off, he sees the police arrive; quickly he parks his car at a nearby park and because he does not want to get caught, he jumps in a moat in the park to hide from the police. Meanwhile, the police have been informed by a witness who saw that a suspicious character got into a red car near the supermarket and drive towards the park with great speed. They find the car just outside the park and search the park, finding Haaknat, who is soaking wet from the water in the moat.

In this story only one causal relation is explicitly mentioned, namely "Haaknat needs money *so* he decides to rob the supermarket". Most of the time, implicit causal relations are somehow assumed by the reader of the story through various *cues to causality*.These cues to causality, which were first mentioned by Hume (1888; see also Einhorn and Hogarth, 1986), are indications that there is a causal relation between states or events. For example, the temporal ordering of the events in a story might tell us something about the causal relations in a story, as events that happen quickly after one another might have influenced each other and an event can obviously not be caused by another event that is later in time. However, we should be careful not to commit the logical fallacy of post hoc *ergo propter hoc* (literally "after this, therefore because of this"). In addition to temporal contiguity, there are other cues to causality such as, for example, spatial contiguity and correlation (Einhorn and Hogarth, 1986). Most of these cues are not a guarantee that there is a causal relation[9] so care should be taken not to assume a causal relation for every (supposed) cue to causality.

The causal relations between the events in a story can be made explicit by expressing them as conditional statements. Figure 3.16 shows the causal relations between the various events in the Haaknat story. For simplicity's sake, not all the individual events and causal relations have been rendered in the figure. Some of the causal links denote not much more than temporal precedence and there are many more implicit causal relations; for example, the fact that the police are searching the park is caused by the event that the supermarket was robbed and that they have an indication the robber is in the park.

In the above story, several types of causal relations can be distinguished (see Warren et al., 1979; Trabasso and Sperry, 1985; Schank and Abelson, 1977, pp. 30–32). *Psychological* or *motivational* causation is assumed to be between physical and mental events and states of affairs: a certain event causes a psychological reaction and a goal or state of mind can motivate someone to perform an act. For example, Haaknat not having any drugs causes him to want some money to buy drugs, which in turn causes him to form the goal to rob the supermarket. *Physical causation* is between two physical states or events. For example, Haaknat jumping in a moat full of water causes him to get wet. *Enablement* involves states or events which are necessary but not sufficient to cause other states and events. For instance, the fact that there is a moat in the park enables Haaknat to hide in it; however, the event that Haaknat jumps into the moat is not caused by the presence of the moat but by his fear of getting caught.

The causal relations between the various states and events in a story can be expressed as conditional statements and these conditional statements are essentially *causal generalizations*. For example, in the Haaknat story, the relation expressed by the generalization "if Haaknat threatens the owner with a gun (demanding money),

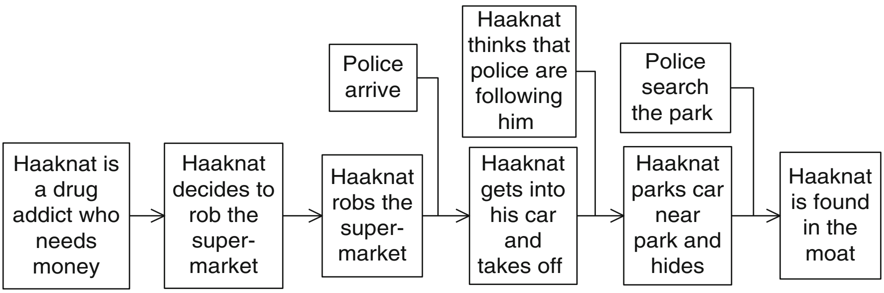



**Fig. 3.16** Causal structure of the Haaknat story

[9] Another well-known fallacy is cum hoc ergo propter hoc or correlation does not imply causation.

then the owner will hand over the money" connects the events that Haaknat threatened the owner with a gun and that the owner handed Haaknat the money. Notice that here, a case-specific generalization has been given, which applies only in the current circumstances (i.e. the circumstances surrounding Haaknat and the owner). Because the causal relations in a story will often be implicit, we are essentially "free to choose" the level of specificity of generalizations that express the causal relations. The abov-mentioned relation, for example, can also be expressed as "if someone threatens another person with a gun (demanding money), then this other person will hand over the money". However, these non case-specific generalizations may often seem implausible. For example, the causal generalization "if someone needs money, he will decide to rob a supermarket" is perhaps somewhat far-fetched, since not many people who need money will rob a supermarket.

Just as with other generalizations, causal generalizations may be supported by empirical findings; the (somewhat more specific) generalization "if person *x* is a drug addict with no money then person *x* will rob a supermarket" may be based on findings by the police that often drug addicts are driven into a life of crime by their addiction. The even more specific "if Haaknat needs money, he will decide to rob the nearby supermarket" may be based on evidence of Haaknat's character or the fact that this is not the first time that Haaknat has robbed a nearby supermarket. Exactly how generalizations can be supported will be further discussed in Chapter 4 on the hybrid theory.

Causal generalizations can be refined in the same way as evidential generalizations. Firstly, a causal generalization can be made more specific by explicitly adding the "hidden conditions". For example, say that we have a generalization "if someone wants to rob another person, this may cause him to threaten that person with a gun". A precondition for this causal inference is that the person who threatens with a gun actually has a gun. Thus the generalization is changed into "if someone wants to rob another person *and he has a gun*, this may cause him to threaten that person with the gun" (Fig. 3.17).

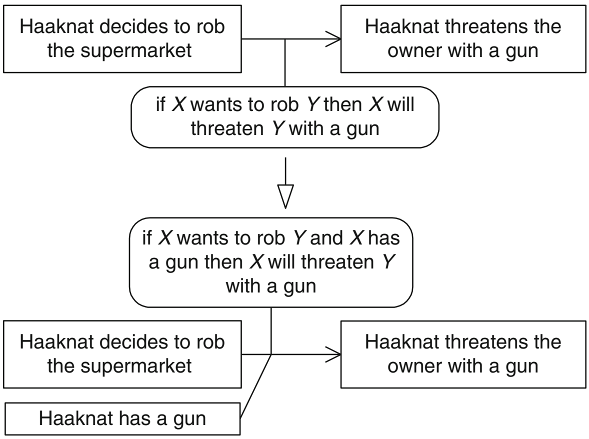



**Fig. 3.17** Adding hidden conditions to a generalization

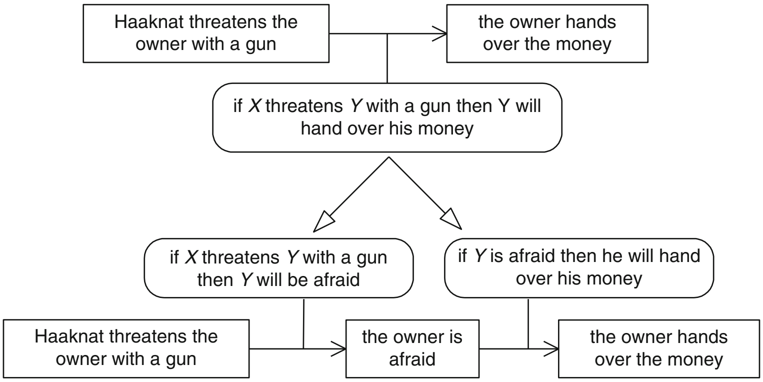



**Fig. 3.18** Unpacking a causal generalization

Secondly, a causal generalization can be changed into a more specific generalization by "unpacking" it, changing one causal link into a causal chain (Fig. 3.18, p. 63). Making the individual causal generalizations in a story explicit and refining them allows one to expose points of doubt in a story in the same way as expressing and refining the evidential generalizations in an argument allows one to identify points of doubt in an argument. If events in a story cannot be causally connected with a sensible causal relation, the story is less coherent and hence less plausible. By refining the causal generalizations, the story is made more specific and its causal connectedness can be analysed.

Wagenaar et al. (1993; Crombag et al., 1994) argue that a story can be analysed by making the causal generalizations in the story "safer", that is, more detailed or refined. Just as with other generalizations, accepting a particular causal generalization depends on one's stock of knowledge and the cognitive consensus about this stock of knowledge. Specifying the causal generalizations in a story allows them to be the subject of debate and therefore may help in finding a cognitive consensus about the coherence of a particular story. Furthermore, it allows for a more refined look on contradictions in a story. A story does not only contradict itself if there are two obviously contradicting states of affairs in the story, but it might also be the case that the causal relations point to a contradiction in the story; for example, when an event *a* is later in time than event *b* but event *b* is supposed to cause event *a*. Exactly how the individual causal generalizations can be debated will be discussed in Chapter 4.

### *3.2.2 Episodic Structures and Story Schemes*

In addition to assuming causal relations through cues to causation, a reader can usually also judge the coherence of a story because a typical story conforms to a certain higher order structure which tells us what types of states and events should be present in such a story. One such structure is the episodic structure developed mainly

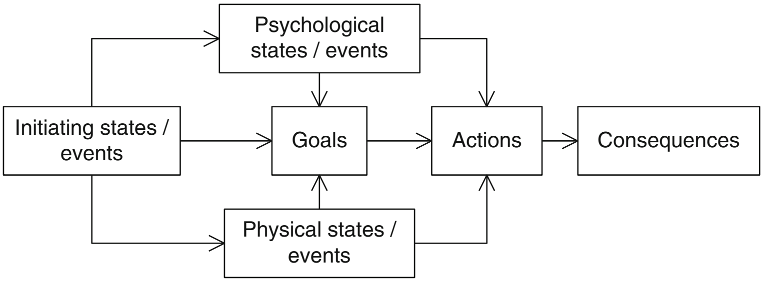


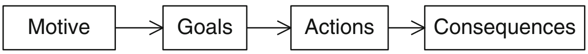

**Fig. 3.19** An episode scheme for intentional actions

in work on story grammars (Rumelhart, 1975; Johnson and Mandler, 1980).These story grammars divide stories into episodes and the events in an episode are categorized according to the role they fulfil in the story. In an episode, *initiating states and/or events* cause some *physical event* and a *psychological reaction* by the actor on the basis of which the actor forms certain *goals*, which in turn lead to *actions* that have *consequences*. In the Haaknat story, Haaknat's addiction (initiating state) causes him to want money (psychological state), so he decides to rob the supermarket owner (goal), which he subsequently does (action), leaving him with the owner's money (consequence). The categories in an episode can be linked by causal relations which are assumed to be present in a typical story (Fig. 3.19).

This specific scheme has been adapted from Pennington and Hastie but it can be seen as expressing the general episode structure as developed in the work on story grammars. Bennett and Feldman's work on legal reasoning with stories also speaks of a central action and a purpose (i.e. a motive or goal) which underlies the action. In the above scheme for intentional actions, the combination of the initiating states and events, physical states and psychological states and events can be seen as a motive in the broad sense (see Section 2.3.3 on motive) and the psychological states and events are a motive in the narrow sense, that is, a specific emotion. Therefore, the scheme can be simplified accordingly (Fig. 3.20). In the rest of these book, both schemes from Figs. 3.19 and 3.20 will be used. Notice that the schemes contain causal relations; in a sense, an episode scheme imposes a causal structure upon a story that is structured according to the scheme. For example, the causal link between "action" and "consequence" indicates there is a causal relation between Haaknat threatening the owner and the owner handing over the money, even though this causal relation is not explicitly mentioned in the story.

More complex stories usually consist of successive episodes, where the events in one episode are the initiating events of the next episode. In the Haaknat example, the episode detailing the events during the robbery and the arrival of the police

**Fig. 3.20** A simplified scheme for motivated actions



(initiating states and events) cause Haaknat to have a psychological reaction (he does not want to get caught) so he takes off (action). Another way for episodes in a story to interact is through goal embedding: an actor forms a subgoal and tries to realize this subgoal so that the main goal can subsequently be realized. In the example, Haaknat's main goal is probably to buy some drugs. Because he does not have the money for this, he forms his subgoal: rob the supermarket. If the robbery is successful, he can achieve his main goal, buying drugs. In Fig. 3.21, the structure of successive episodes and sub episodes has been represented. Here, $I_e$ are the initiating states of episode *e*, $P_e$ its psychological states, $G_e$ its goals, $A_e$ its actions and $C_e$ its consequences. These structures have been adapted from (Trabasso and van den Broek 1985), who give some more examples of how episodes can interact and form complex structures. However, the two structures in Fig. 3.21 are sufficient for current purposes. Notice that episode such as the ones shown here are *hierarchical* in that, for example, the *initiating states and events* from the episode in Fig. 3.19 can itself also be an episode. This is shown in Fig. 3.21, where episode $e_1$ forms the initiating states and events for episode $e_2$.

The episodes as discussed below essentially provide a general structure for intentional actions. If a story conforms to this structure, it is judged as more coherent because the states and events in the story are structured in such a way that we expect things to happen in the world around us. If the story is structured according to the above episode structure, then the story contains clear motivations for why an actor in the story has particular goals and we know that the actor's actions are driven by goals and that the consequences of the actions are made explicit. For example, if we were to leave out the information that Haaknat is a drug addict and that he needs money, then it would not be clear exactly why Haaknat robs the supermarket. This makes the story less coherent, because we do not expect normal people to

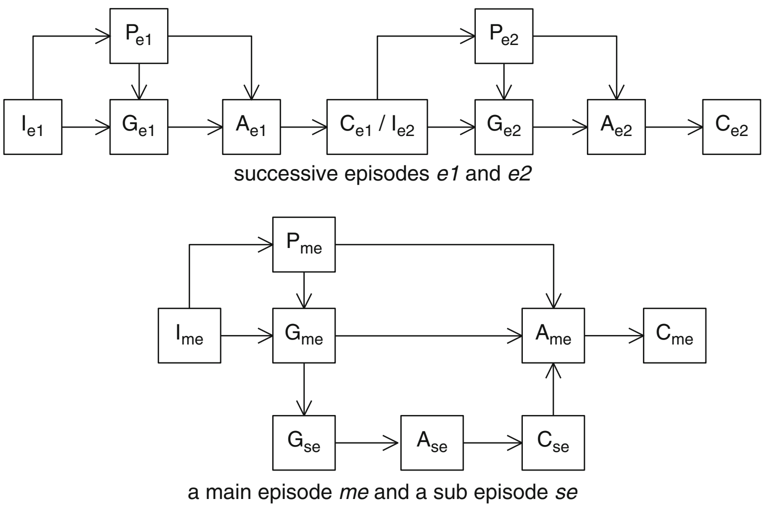



**Fig. 3.21** Complex episode structures

rob a supermarket, that is, we like to see that someone has a good motivation for robbing the supermarket. Furthermore, encountering different types of events (i.e. goals, actions, consequences) in a story can also be seen as a cue to causality: for example, we know that (physical) actions usually cause some kind of consequences in the physical world. Thus the higher order episode structure suggests (implicit) causal relations between the events in the various categories.

Further research has shown that, in addition to the general intentional episode scheme of "motive – goal – action – consequence" more context-dependent information about standard patterns of actions can also be identified. Detailed scripts (Schank and Abelson, 1977) or explanation patterns (Schank, 1986) also provide the structure of a typical story, but instead of structuring the story along general categories like "action" and "goals", a script provides a prototypical story for a certain specific situation. The much quoted "restaurant script", for example, contains information about the standard sequence of events that take place when somebody goes to dine in a restaurant. Scripts help us to understand stories by filling in missing information. As an example, take the following (very short) story:

> Henry went to a restaurant. He asked the waitress for a plate of spaghetti. He paid the check and left.

This story is understandable because it references to the restaurant script. Not all of the details (Henry taking off his coat, Henry reading the menu etc.) have to be mentioned because they are part of the implicit restaurant script we have memorized through our experience with restaurants. Scripts contain not just a pattern of actions, but also define the roles of different actors and provide additional information. Take, for example, a very basic version of the script for "robbery":

> **Roles:** robber ($x$), person being robbed ($y$)
> **Relevant information:** the motive for the robbery $m$, the time of the robbery $t$, the place of the robbery $p$, the type of force employed $f$, the nature of the goods $g$
> **Pattern of actions:** $x$ has a motive $m$ for wanting $g$ – $y$ owns $g$ – $x$ wants $g$ – $x$ wants to rob $y$ of $g$ – $x$ has an opportunity to rob $y$ – $x$ robs $y$ – $y$ loses $g$
> **More specific kinds of robbery:** armed robbery, mugging, carjacking

The structure of the above script is based on Schank and Abelson's work on scripts and Schank's work on explanation patterns. In addition to a general pattern of actions we expect to come across in a story about a robbery, other information which might be relevant in a story about a robbery is also given. This additional information can also help make sense of a story and making this information explicit improves the coherence of a story. For example, a story about a robbery which does not mention what was stolen is slightly puzzling. Following Schank, more specific kinds of the general script are also given. These specific kinds of robbery also have their own scripts with more detailed information. For example, the script for "armed robbery" will also have an element "type of weapon $W$ used". Like episode schemes, scripts determine what we expect in a story and hence when a story can be judged as complete. For example, a story about an armed robbery which does not mention when, where and with what weapon the robbery was committed is less complete and arguably less plausible than a story that explicitly mentions all these elements.

Both episode schemes and scripts can be seen as instances of something which I will call *story schemes* (Bex, 2009).[10] These schemes divide the various states and events in the story into different categories, ranging from abstract (e.g. "actions") to more specific (e.g. "*x* robs *y*"), and sometimes define the temporal and causal relations between these categories.An important parallel can be drawn between generalizations and story schemes. Where generalizations can often be seen as conditionals that act as a general background for a single inference, story schemes are more complex structures that act as a general background for some story.[11] If we compare story schemes with argumentation schemes, we could say that an argumentation scheme is a general scheme for arguments of a particular kind just as a story scheme is a general scheme for stories of a particular kind.

Both generalizations and story schemes are part of our stock of knowledge and express general knowledge about the world around us. Like generalizations, story schemes and stories play an important part in "making sense" of the world because they order information about human actions in a format that is more easily understandable than complex conditional generalizations; Schank and Abelson argue that our memories are often organized as stories or general scripts. So story schemes and stories are just as necessary as generalizations when performing commonsense reasoning.

However, like generalizations, story schemes can also be dangerous in that they can express false beliefs, prejudices, ideal but non-realistic situations etcetera. Recall that one of the main dangers of generalizations was that they are often left implicit. The same is true for story schemes (maybe even more so). Because this knowledge is left implicit, we risk falsely inferring what we think is general knowledge (i.e. story schemes) from particular examples (i.e. stories); as was argued above, like generalizations many story schemes do not follow from empirical research but are "synthesized" from our stock of knowledge. Here, the reliability and commonality of a story scheme are important. That people usually act according to some psychological motivation has been empirically tested and is commonly accepted. Furthermore, that actions often have certain consequences is also something not many people would deny. Hence, the intentional action scheme can be seen as quite common and reliable. However, because the abundance of stories we are confronted with in everyday life, we might base our supposed general knowledge on bad examples. For example, from watching shows and films on TV I might think that I have correct knowledge about what the average American court session looks like: the (usually innocent) defendant is assisted by a heroic young lawyer who tells moving stories about prejudices against people of the defendant's age, race or gender, while the bitter and straight prosecutor yells "objection!" all the time. In the end, the jurors decide with their heart and the defendant is acquitted. However, this is probably not how a typical court case takes place, at least not outside of a Hollywood studio.

---

[10] Twining (1999) calls these schemes "story types" or "scenarios".

[11] In this sense, a particular story can be likened to a case-specific generalization.

The dangers associated with stories do not mean that they cannot or should not be used in reasoning with evidence. As in reasoning with arguments and generalizations, the dangers of stories and story schemes can be lessened by specifying exactly which schemes we use, how we use these schemes and from which sources the schemes stem. Making the knowledge from our shared stock explicit in this way minimizes the chance that it is misunderstood or used in the wrong way. Furthermore, sources of doubt related to a particular story scheme can be identified just as they can be identified for a generalization: *critical questions* can be associated with a story scheme in the same way as critical questions are associated with an argumentation scheme. For example, the following critical questions concern the intentional action scheme:

- Is the psychological state a proper motive for the goal and the subsequent action?
- Can the action be performed given the initial and physical states?
- Does the action have the stated consequences?

These critical questions are adapted from (Bex et al., 2009), who use them for the (abductive) argumentation scheme for practical reasoning (see Section 3.1.3).

Like generalizations, story scheme also vary from abstract to specific. The more specific schemes can be seen as instances of the more abstract schemes; for example, a robbery is an instance of an intentional action. In this way,story schemes can be said to *correspond to* each other: a story scheme $S_1$ corresponds to a story scheme $S_2$ to the extent that the elements in $S_1$ correspond to the elements in $S_2$. The relations between the elements of the intentional action and the robbery scheme are shown in Fig. 3.22. These *abstraction relations* between the elements of the schemes can be expressed as generalizations. For example, "*x robs y* is an *action*". These generalizations have to be plausible: when such a generalization is implausible, the two schemes do not match. For example, the abstraction generalization "*y owns g* is a *goal*" is not plausible so there is no abstraction relation between these elements. Sometimes two matching schemes only differ on one or two elements. For example, the element "*x* robs *y*" in the robbery scheme can also be modelled as two separate elements, "*x* threatens to use force against *y*" and "*y* hands over *g* to *x*", thus making the concept of "robbery" more specific. The more specific types of robbery mentioned above (armed robbery, carjacking) also differ from the general robbery scheme on only a few elements (the type of force used, the nature of the goods).

It is also possible that a particular story corresponds to a story scheme. Sometimes this correspondence involves simply instantiating the variables in the scheme

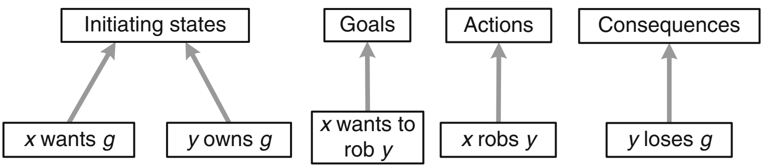



**Fig. 3.22** Correspondence between the intentional action and the robbery scheme

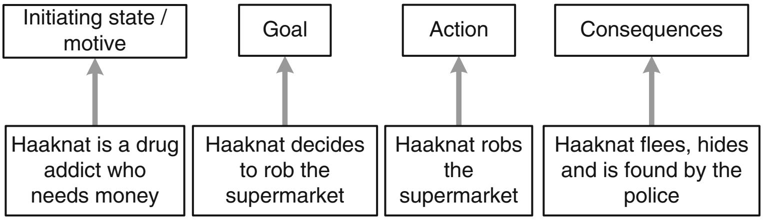



**Fig. 3.23** Correspondence between the intentional action scheme and the Haaknat story

with particular terms. For example, the Haaknat story from p. 62 is essentially a particular instantiation of the robbery scheme. The Haaknat story also roughly corresponds to the intentional action scheme through abstraction relations, viz. Fig. 3.23. While it has not been explicitly shown in the above figures, the causal structure of a story and a scheme should also correspond for the story to properly match the scheme. If, for example, in the Haaknat story Haaknat hides *before* he robs the supermarket, then his hiding is clearly not a consequence of the action that he robs the supermarket.

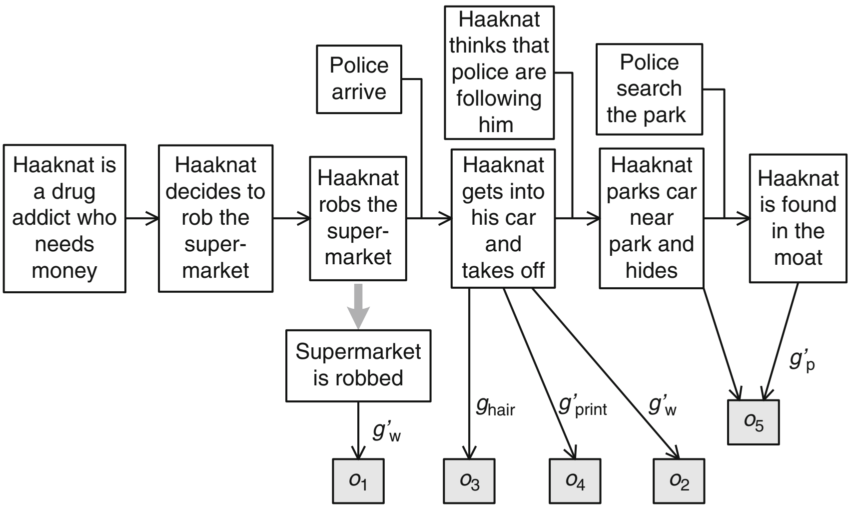
In addition to having features in common, story schemes and generalizations are also closely related in other ways. The most obvious relation is that causal generalizations can be part of a story scheme. In a way, a story scheme can be seen as a chain of generalizations and a complex generalization can be rephrased as a story scheme. Take as an example the complex generalization from Fig. 3.15; recall that this generalization was somewhat contrived, viz.: "if a supermarket is robbed and there is a man who has a clear motive for robbing a supermarket and this man could have been the person robbing the supermarket (based on his whereabouts just after the robbery) and this man was acting suspiciously when apprehended then this man is the person who robbed the supermarket". Now, this generalization can be rephrased as a simple story scheme: $x$ has a motive for robbing $y$ – $x$ has the opportunity to rob $y$ – $x$ robs $y$ – $x$ flees – $x$ acts suspicious – $x$ is apprehended. This story scheme is essentially a slightly different version of the above robbery scheme with a "fleeing" episode attached.

While specific causal relations can be part of a story scheme, they are often not rendered as a detailed causal structure and they are not concerned with all the possible implicit causal relations. Rather, a story scheme should be viewed as a more "holistic" knowledge structure which denotes the general structure of a story. In this general structure, it is not the individual causal generalizations that matter but rather their interaction in the scheme.

### *3.2.3 Explaining the Evidence*

In the story-based approach to reasoning with evidence, the main goal is to construct a hypothetical story, which represents and makes sense of "what happened" in a case. This story should properly *cover* the evidential data by *causally explaining* this data.

Pennington and Hastie do not discuss in detail exactly how a story causally explains the evidential data. However, the idea of a causally connected story being an explanation for the evidential data was elaborated on by researchers in AI. Formal models of causal-abductive reasoning were proposed by, for example, Console and Torasso (1991, see Lucas, 1997 for an overview) and researchers such as Josephson (2002) and Thagard (2004) elaborated on this basic approach and applied it to reasoning with criminal evidence. They argue that reasoning with stories takes the form of abductive inference to the best explanation, where the story or stories are abductively inferred from the evidence that has been observed. Such abductive reasoning can be modelled in various ways (see Section 2.3.2 on causal and abductive reasoning in AI) but basically the observed pieces of evidence are the *observed effects* or *observations* that causally follow from the story: if we have a causal relation expressed by the generalization "if *cause* then *evidence*" and we observe *evidence*, we are allowed to infer *cause* as a possible explanation of the evidence. This cause which is used to explain the effect can be a single state or event, but it can also be a sequence of events, a story. The Haaknat story from p. 62 and Fig. 3.16 can now be modelled as a simple causal model that explains the observed evidence in the case as in Fig. 3.24.




$g'_w$: if *p* then a witness testifies that '*p*'

$g_{hair}$: if a person *x* gets into a car this will cause *x*'s hair to be in the car

$g'_{print}$: if person *x* gets into a car this will cause *x*'s shoeprint to be in the car

$g'_p$: if *p* then the police report will state '*p*'

$o_1$: Witness $W_2$: 'I saw someone with a gun rob the owner'.

$o_2$: Witness $W_1$: 'I saw someone looking like Haaknat get into the car'.

$o_3$: Haaknat's hair was in the car.

$o_4$: The shoeprints found in the car were of the type of shoes Haaknat wore when apprehended

$o_5$: Police report: 'The car was found near the park and Haaknat was found in the moat'.

**Fig. 3.24** Observed evidence explained by the story

The arrow between the events that "Haaknat robs the supermarket" and "Supermarket is robbed" is not a causal relation but an abstraction relation (see the second part of Section 3.2.2): that Haaknat robbed the supermarket is an instantiation of the more abstract event that the supermarket is robbed. This abstraction relation can thus not only be used to connect a story to a more abstract scheme but also to connect one or more events in a story to a more abstract version of the event. This allows for the possibility of *hierarchical stories* with multiple levels of abstraction.[12] In Section 5.3.3 this way of using abstraction generalizations will be further explained in a formal setting.

In Fig. 3.24 each observation follows from an event in the story and the story can thus be said to causally explain the evidential data in the Haaknat case. Notice that not all the evidential data is explicitly mentioned; in the case of the shoeprints and matching hairs the story only explains propositions that in the argumentative approach were said to follow from the evidential data (see Fig. 3.4). In other words, it is not shown exactly how the story explains the expert testimony or how the fact that the hairs match follows from the expert testimony. In the story-based approach it is often only roughly outlined how exactly the data that is explained is related to the evidential data in the case. This fits in with the holistic and story-centred idea of this approach: the focus should not be on the details of exactly how individual pieces of evidential data follow from the story but on the global overview. A story gives a quick and easily understandable overview of the case and the evidential data and such a summary allows us to organize the evidential data as well as our own thoughts in a case. By looking at which observations are not yet covered, we can judge which parts of the case should be further analysed.

In the causal-abductive approach to reasoning with stories as specified here it is, however, possible to perform a more detailed analysis of how the evidential data causally follows from the story. For example, the Haaknat story from Fig. 3.24 can be an explanation for the expert testimony as shown in Fig. 3.25. Notice that the

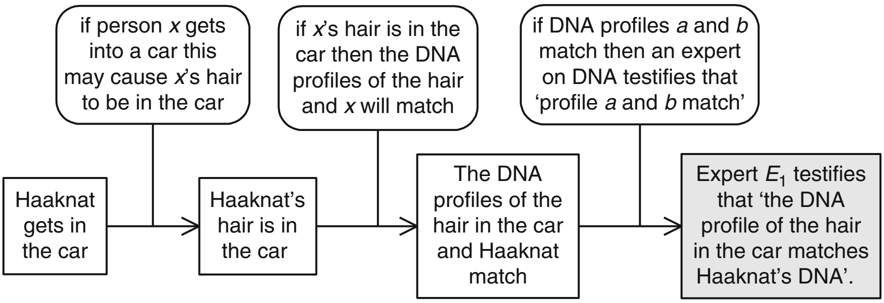



**Fig. 3.25** An event explaining the expert testimony

[12] Just as a story can correspond to an abstract story scheme, a story can also correspond to a more abstract (but still case-specific) or more specific version of itself. For example, the event "Haaknat robs the supermarket" can be specified as "Haaknat threatens to use force against the owner of the supermarket" and "the owner of the supermarket hands over the money to Haaknat".

causal reasoning from "Haaknat was in the car" to the expert evidence is essentially the inverse of the subargument about the expert testimony in Fig. 3.4. Such detailed causal structures which model how evidence follows from an event in the story also allows for defeasible reasoning about individual pieces of evidence. Abductive inference from data to a hypothesis is defeasible in that there may be other explanations for the data. For example, in Fig. 3.23 witness $W_1$'s testimony may be explained by the hypothesis that Haaknat got in the car while the witness was watching, but it can also be explained by the hypothesis that someone else who looks like Haaknat got in the car while the witness was watching. The witness may also have another reason which causes him to testify that he saw someone looking like Haaknat; maybe he misremembers or he is lying (Fig. 3.26).

In this situation, a choice will have to be made what the most likely explanation of the three is. The underlying idea of the story-based approach, however, is not that the individual pieces of evidence are scrutinized in this way. Rather, the focus is on how much of the total set of evidential data a story explains (Pennington and Hastie call this the *evidential coverage* of a story). The idea is that ultimately, the evidence which for some reason should not be believed or trusted does not follow from the best story that explains most of the evidence. In the case of the Haaknat story, there is only one proper story[13] that explains that the witness saw Haaknat, namely the story from Fig. 3.24. If we want to provide an alternative for this testimony (e.g. one of the alternatives from Fig. 3.26), this alternative will have to be incorporated in a proper story. For example, the proposition "the witness has some other reason to testify that he saw someone looking like Haaknat" does on its own not provide a proper alternative to the Haaknat story from Fig. 3.24. If, however, it were to be incorporated in a proper story that the witness wants to frame Haaknat because of some past argument with Haaknat, it would constitute a valid alternative explanation for the witness testimony evidential data.

In addition to the extent to which a story explains evidence, it is also important to consider a story's *consistency* with the evidence, whether or not it contradicts evidential data. Say that, for example, a close friend of Haaknat's testifies that Haaknat

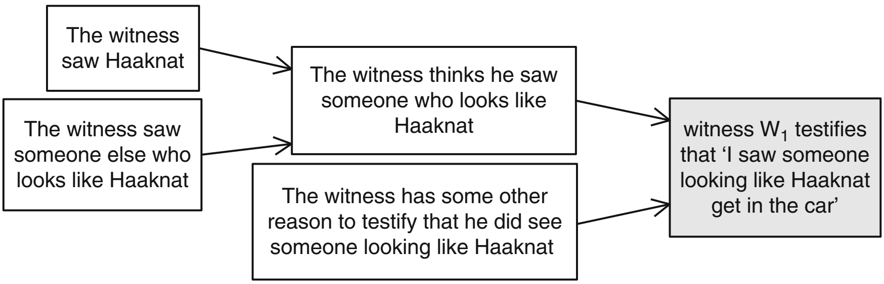



**Fig. 3.26** Three explanations for a witness testimony

[13] Exactly what constitutes a proper or a good story will be discussed in Section 3.2.4.

has been off his drug habit for the last 4 years and working a steady job. This directly contradicts the story in Fig. 3.24. This contradiction can also be less direct. Say, for example, that we have a new witness who testifies that he saw Haaknat was pushed in the moat by another man who was running through the park. This contradicts that Haaknat jumped into the moat himself. Finally, it is also possible that evidential data contradicts something that is not in the story itself but which follows from the story through predictive reasoning. For example, from the event in the story "the owner hands Haaknat the money" it follows that not all the money is in the supermarket after Haaknat left. Now, if the police find that all the money is still in the supermarket when they arrive, this clearly contradicts the story that Haaknat took the money with him. Like evidential coverage, contradiction is again not precisely defined by Pennington and Hastie. We can say that a story contradicts some event if it explains the negation of the event. However, a piece of evidence itself cannot be negated; only an event inferred from a piece of evidence can be negated. In Chapter 4 this will be further discussed.

### *3.2.4 Choosing the Best Explanatory Story for the Evidence*

In the story-based approach, alternative stories that denote possible hypotheses of what happened in the case are constructed. In the story-based approach, a story should not only be compatible with the available evidential data but it should also be coherent. Furthermore, stories should also be compared to any possible alternative stories in order to determine what the best story is. This comparison of alternative stories about what happened is thus a process of *inference to the best explanation*.

Pennington and Hastie discuss in some detail the various criteria that determine the quality of a story and thus the confidence people have in a story. They argue that a good story (i.e. a story in which the confidence is high) should *explain* the available evidence, it should be *consistent* with the available evidence and it should be *coherent*. These criteria can also be applied normatively in the process of proof to determine which of the hypothetical stories in the case is the best.

As an example, take the criteria that pertain to a story's conformance to the evidence: a story should explain the evidence and it should be consistent with the evidence. If a particular story explains more evidence than another story and is contradicted by less evidence, we would say that it better conforms to the evidence and that it is therefore a better story. This corresponds to ideas from research on inference to the best explanation (see Section 2.3) that a hypothesis should explain as much of the data as possible. Take, as an example, the Haaknat story. Haaknat himself gave another explanation for why he was in the moat:

> On the 21st of October, an hour before he was found in the moat, Haaknat had an appointment with Bennie, who owed him some money. Bennie did not want to pay back the money and Haaknat and Bennie got into an argument. Haaknat then ran towards the park. In the park, Haaknat saw the police officers looking for someone and he jumped into the moat to hide. Some time later, the police found Haaknat in the moat.

While this story explains the fact that Haaknat was found in the moat, it does not explain the robbery and it is also not explained when and where Haaknat drove in the car. While it seems that this story should not even be considered because it does not explain the robbery, it should be noted that evidential coverage is a relative concept in that we cannot always expect a single story to explain all the evidence. In any case, there are several explananda and it is possible that there are independent explanations for different explananda. In the above example, we might say that if we believe Haaknat's alibi story it does not need to explain the robbery and that in that case the robbery should be explained by another, independent story which has nothing to do with Haaknat. Effectively we are comparing the original story that argues that the explananda (that the supermarket was robbed and that Haaknat was hiding in the moat) are somehow linked with two other stories that each provide an independent explanation for an explanandum and thus argue that the robbery and Haaknat hiding in the moat are really independent occurrences. In this example, however, I will for simplicity's sake only compare the original story with the second story that explains why Haaknat was in the moat. While we have to be cautious in writing off a story because it does not explain one of the explananda, in this case we can say that the original story better covers the evidence because it also explains Haaknat's hair and shoeprints in the car while the second story does not explain this. None of the two stories is contradicted by evidential data.

It is possible that two explanations have the same coverage and that one of them explains the evidence better than the other. This happens when the causal connection between one explanation and the evidence is stronger than the causal connection between the other explanation and the evidence. For example, say that Haaknat explained the shoeprints and the hair by arguing that someone who had the same shoes as him brushed past him and that thus one of his hairs stuck to this other person's coat and that subsequently this person must have gotten into the car, leaving behind the shoeprints and the hair. This explains the shoeprint and the hair, but in a more convoluted way than the original story and in this case, we might argue that the causal connection between the story in Fig. 3.24 and the evidence is stronger than the causal connection between Haaknat's explanation (that he brushed past someone).

The quality of a story depends not only on the extent to which it conforms to the evidence, but also on a story's coherence or plausibility, that is, whether or not the story conforms to the general knowledge about the world around us. Josephson and Josephson (1994) argue that "we should be cautious about accepting a hypothesis, even if it is clearly the best one we have, if it is not sufficiently plausible in itself". The coherence of a story can essentially be determined irrespective of the evidential data. In fact, Bennett and Feldman (1981) have shown that people's confidence in a story often depends not on whether it actually happened, but rather on the coherence of the story. In an experiment, they asked 85 students to assess the truth of stories that were told by other students. Some of these stories were really true (that is, the events recounted had really happened) and other stories were made-up. Some of the stories (both true stories and made-up stories) were coherent but other stories were incoherent. It turned out that there is a significant relation between the structural coherence of a story and its credibility. That is, the more coherent a story, the higher the probability that the story is judged true, irrespective of the *actual* truth of the story.

In the above Sections 3.2.1 and 3.2.2, two factors that are important for a story's coherence were discussed, namely its causal connectedness and the extent to which it conforms to a story scheme.Pennington and Hastie argue that a story's coherence depends on three sub criteria: consistency, plausibility and completeness. A story should be *internally consistent* in that it does not contain internal contradictions between different parts of the story. A storyis *plausible* if it conforms to the decision maker's general knowledge of the world and a story is *complete* when all of the elements of the episode scheme for intentional action (Fig. 3.19, p. 64) are part of the story. Pennington and Hastie have adapted these criteria from the general criteria of story coherence as discussed in the above Sections 3.2.1 and 3.2.2.

The first criterion that is used by Pennington and Hastie to determine the coherence of a story, which was already briefly mentioned at the beginning and end of Section 3.2.1, is that of *internal consistency*. This means that the stories should not contain two states or events that clearly contradict each other and that a story should also not be contradicted by states or events which follow from the story through predictive reasoning. For example, a story in which Haaknat boards an airplane to New York on the 12th of January at 12:00 and robs a supermarket that same day at 13:00 obviously contradicts itself.

The criterion of *plausibility* is not further explained by Pennington and Hastie. One of the ways to interpret this criterion is to look at the plausibility of the (implicit) causal generalizations and events in a story. For example, a story that talks of little green men from Mars robbing a supermarket is not very plausible. The same goes for a story in which a man was killed because he was shot with a water gun. In addition to the inherent plausibility of the events in the story, an important question is whether all the states and events in the story can be connected by plausible causal relations, that is, relations which are expressed by plausible generalizations. For example, in the story on p. 77, it is not exactly clear why Haaknat flees. That is, it is not clear what the causal connection between the beginning of the story and Haaknat and Bennie's argument and the fact that Haaknat suddenly runs away is. We could expand the story by saying that Haaknat ran away because he felt threatened by Bennie (Fig. 3.27). Notice that here the refinement operation from Fig. 3.17 is performed. By explicitly incorporating this new causal relation into the new story it becomes more plausible: after all, people who are threatened often run away.

The final sub criterion of coherence Pennington and Hastie mention is *completeness*. The basic idea is that the story completes a plausible story scheme, that is, that all the elements of the story scheme correspond to some event in the

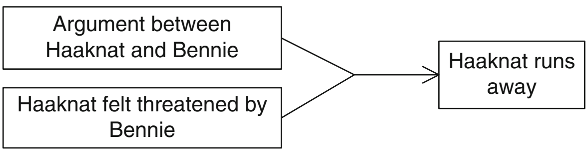



**Fig. 3.27** Refining Haaknat's story

story. Pennington and Hastie argue that the intentional action scheme is the general scheme to which all stories should conform. If a story completes this scheme, it is ensured that the story has "all its parts", i.e. some clear motivating events that initiate a psychological reaction on the basis of which goals are formed which lead to an action that has consequences. In Fig. 3.23 it was shown that the original Haaknat story completes the intentional action scheme. Haaknat's own story, however, story is less coherent than the original one on this point, as there is no clear cause for Haaknat hiding in the moat (Fig. 3.28). Note that in this figure the correspondence relations between the event in the story and the elements of the scheme have not been explicitly represented. Here no clear psychological motive or goal is provided so it seems that Haaknat does not behave in the way we expect people to behave, as normal people do not jump into moats without a good reason (i.e. a motive). In order for a story to correctly complete a scheme, it is required not only that all the elements of the scheme correspond to events in the story, but also that the abstraction relations between the events and these elements are plausible. It is therefore not possible to match one of the initiating events from Fig. 3.28, for example "Haaknat runs away", to the "psychological state" element of the scheme, as "Haaknat runs away denotes a psychological state" is clearly implausible.

There are now three clear criteria for the coherence of a story: it has to conform to a plausible story scheme, it has to be internally plausible in that its events and internal causal relations should be plausible and the story should be internally consistent. Measuring a story's coherence is less obvious than measuring its evidential coverage; with coverage, the number of pieces of evidence can be counted and thus a crude measure of coverage can be defined. Such a measure is harder to provide for coherence. Assessing the internal causal relations in a story involves first making them explicit and the number of plausible generalizations hence depends on how many of these relations have been expressed by explicit generalizations. Furthermore, coherence also involves determining what is a "plausible generalization" or a "plausible story scheme" and this is relative to the cognitive consensus on a stock of knowledge. While a clear measure of a story's coherence cannot be given, the most important idea here is that the story schemes and generalizations on which express the world knowledge used in a story are made explicit, so that possible sources of doubt or implausible schemes or generalizations can be clearly identified. This will be further discussed in Sections 4.3 and 4.4.

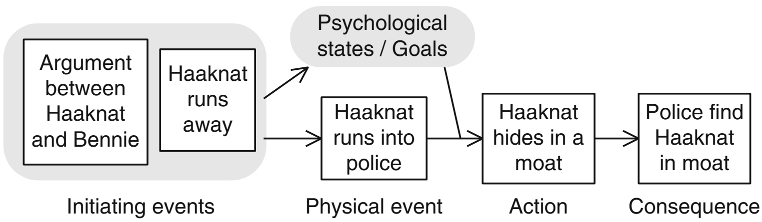



**Fig. 3.28** Trying to match a story to a scheme

In addition to the extent to which a story conforms to the evidence and its coherence, we should also look at how many other possible explanations there are and at how thorough the search for alternative hypotheses has been (Josephson and Josephson, 1994). As for the number of alternative explanations, this should be viewed as pertaining to all explananda. In the Haaknat example, recall that the alternative story only explained some of the explananda. Thus the original story can be judged as "unique" when it comes to explaining the explanandum that there was a robbery.Pennington and Hastie argue that if there are multiple coherent explanations for the evidence, the confidence in each of these stories is decreased. This does not mean that the search for new stories which explain the evidence should be stopped once there is one coherent story. A problem in criminal investigation and decision making is "tunnel-vision" or confirmation bias (see Wagenaar et al., 1993), where one explanation is taken as the right one and the investigation focuses on finding evidence that supports this explanation while dismissing evidence that contradicts this explanation. Determining whether the search for alternative hypotheses has been sufficiently thorough depends not only on rational grounds but also on pragmatic considerations. For example Josephson and Josephson (1994) argue that we should look at the costs of being wrong and the benefits of being right, and how strong is the need to come to a particular conclusion.

Summarizing, the extent to which a story conforms to the evidence and the extent to which it is coherent are criteria that can be used to determine the confidence in a story as well as for comparing stories. Here, a story's conformance to the evidence can be considered as the most important criterion and coherence is more a secondary requirement. The reason for this is that a coherent story that does not explain any evidence should of course not be considered a good hypothesis for what happened in a case. The experiments performed by both Bennett and Feldman (1981) and Pennington and Hastie (1986, 1988, 1992) show that The more coherent a story, the higher the probability that the story is judged true, irrespective of the *actual* truth of the story. This can be dangerous: Crombag and colleagues (1994, Chapter 4) warn about a "good" (i.e. coherent) story which does not conform to the evidence pushing out a "bad" (i.e. incoherent) story which does conform to the evidence. In other words, a good story can and often will win over a true story.

### *3.2.5 Summary and Evaluation*

This chapter has introduced stories as a general form of reasoning and communication. It has shown that the coherence of a story depends on its causal connectedness and the extent to which it conforms to a holistic story scheme. It was shown how stories can be used to abductively explain the evidence and how stories can be assessed and compared according to the extent to which they conform to the evidence and their coherence. This section evaluates the use of stories in reasoning with criminal evidence. First, some specific uses of stories in the three contexts of the process of proof (discovery, pursuit and justification) will be briefly discussed. The holistic and causal nature of stories has certain advantages and disadvantages and these will also be discussed below. This section ends with briefly discussing the naturalness and rational well-foundedness of the story-based approach.

Stories and story schemes can be used in the context of discovery to construct scenarios about what happened in the case. Such hypothetical scenarios provide an overview of the possibilities in the case and allow for a first selection of the most plausible scenarios. By performing further predictive reasoning in the discovery and pursuit stage, further observables may be predicted from the hypothetical stories and thus the search for evidence can be guided by the hypothetical stories. For example, if it was busy in the park with runners and families taking a walk, Haaknat must have been seen by someone when he walked from the car to the moat so the police should ask people who were in the park that day if they saw someone. In this way, the combination of explanatory and predictive reasoning can be used to test a hypothesis in the context of pursuit.

The story-based approach organizes the evidence during the process of explaining and testing hypothetical scenarios as it shows which of the evidential data is already explained by one or more hypotheses and which data still needs to be explained. Stories can be tested in the context of pursuit by looking at the extent to which a story conforms to the evidence and its coherence. In the justification phase, stories play an important role in helping people make sense of what happened in the case. They can make a complex case intelligible by clearly identifying a central action, its consequences and the motivation for this action.

The story-based approach is often called a *holistic* approach because the evidential data is not scrutinized individually but rather as a whole. In this way, the story provides the circumstances under which the crime happened and fills in the gaps in our reasoning by placing events for which there is no evidence in a coherent story (Hypothetical) stories provide a good overview of what (might have) happened in the case allow us to easily understand what a particular case is all about. Stories and in particular story schemes provide for a holistic way of reasoning with general commonsense knowledge. Instead of reasoning about a situation with individual (conditional) generalizations, stories and story schemes allow us to construct and consider the situation “as a whole”. By using story schemes in the context of discovery for the quick and relatively easy construction of hypothetical scenarios, the danger of tunnel vision can be lessened.

The most important advantage of this holistic and scenario-focused way of reasoning with evidence is that it is closest to how investigators and legal decision makers actually think about a case.[14] Experiments by Bennett and Feldman (1981) and Pennington and Hastie (1986; 1988, 1992) suggest that when making a decision, jurors construct and compare stories which explain the evidence and then choose the most coherent and plausible story that covers the most evidence. Stories help people organize the evidence and make sense of a case. The causal and schematic structure

[14]Twining (1999), however, argues that stories are not always necessary. In many criminal cases there is only one fact at issue and complex hypotheses involving motives and actions need not be considered.

of stories helps the decision maker in remembering the evidence and judging the importance of evidence. Through their structure, stories also help people to fill gaps in a case. For example, the suspect's intentions must often be inferred from the actions that the suspect performed; in other words, these intentions can be inferred from the story. Studies in story recall and understanding have also shown that our memories are organized through episodes or stories (Schank, 1975, Stein and Glenn, 1979).[15] Furthermore, there seems to be a consensus in the literature on evidence analysis that fact investigators work with causal story structures and timelines (De Poot et al., 2004; Heuer, 1999; Kerstholt and Eikelboom, 2007, see Section 2.2.1). A criminal case which has to be investigated is interpreted through different scenarios or stories which reconstruct what might have happened. These scenarios are then tested and the most likely scenario is chosen. These findings were informally confirmed in our contacts with police detectives and lecturers of the Dutch police academy, in which we learned that crime investigators often visualize time lines and simple scenarios to make sense of a body of evidence. Furthermore, Pardo and Allen (2007) argue that Inference to the Best Explanation using (explanatory) narratives is the right way of modelling reasoning with evidence in the context of both civil and criminal trial.

The holistic approach of stories provides a psychological advantage but it also has some inherent disadvantages and dangers. One of the main dangers of stories is that a coherent story is judged as more believable than an incoherent story, regardless of the actual truth of the story. This danger can be obviated by the requirement that the story should explain as much of the evidential data in the case. In this respect, one of the main disadvantages of the holistic approach of stories is that the evidential data often has no clear and separate place in the model of the case. Purely story-based theories require that the observations are explained by the story, but it is not made clear whether these observations are the actual evidential data itself or whether only the events that follow from the evidential data are the observations or whether the evidential data is in some other way part of the story. In other words, the important distinction between the evidential data $E^*$ and the event $E$ as made in Section 2.1.2 on evidence is often not made in literature that follows mainly a story-based approach (e.g. Bennett and Feldman, 1981; Pennington and Hastie, 1993b; Wagenaar et al., 1993). For example, Pennington and Hastie mention none of the evidential data explicitly. Instead they argue that the elements of the story *are* the evidence, that is, items of evidence appear directly in the story. For example, they mention "Caldwell was in Gleason's Bar" as a piece of evidence appearing directly in the story (Pennington and Hastie, 1993a, p. 131). In reality, the piece of evidence is probably something like "witness $W$ testified that Caldwell was in Gleason's Bar" and the event that Caldwell was in the bar follows from this piece of evidence. However, in their example Pennington and Hastie do not mention such pieces of evidence explicitly and it is not always clear what is part of the story and what is

[15] In this sense, stories play a big *psychological* role (Anderson et al., 2005), as opposed to the more atomistic arguments, which play a more logical role.

considered to be evidential data. Because the evidential data and the inferences made based on this data are not given a clear position in the story-based approach, issues like witness credibility or expert bias are not mentioned. Pennington and Hastie argue that "the evidence item "Caldwell was in Gleason's Bar" is direct testimony, is not a matter of dispute [...and therefore...] this is shown as piece of evidence appearing directly in the story" (Pennington and Hastie, 1993a, p. 131, 132). Note that here it is not made clear who testified that Caldwell was in the bar and why this particular evidence is not a matter of dispute. In sum, the important conceptual ambiguity between a story and the evidence on which it is based is not made in the story-based approach.

Work in AI by Josephson (2002) and Thagard (2004) has elaborated on the story-based approach and proposed a way in which the evidential data can be connected to the story, namely by requiring that the story is a causal explanation for the evidence (see Figs. 3.25, p. 71 and 3.26, p. 72). While this way of modelling the connection between a story and the evidence does not fully disambiguate between what is the story and what is the evidence, modelling the link between stories and evidence in this causal way alleviates the problem that the evidential data does not properly follow from the story. Thus, it allows for reasoning about the individual pieces of evidence: as Fig. 3.26 has shown, it is possible to reason about matters of witness credibility or observational sensitivity in a purely causal approach by comparing the various causes for the testimony. Furthermore, a particular advantage of causal reasoning with stories is that it allows for the prediction of so-called *story-consequences*, possible events which should have happened if we take the story to be true.

However, there are indications that this way of treating simple alternative explanations does not adequately capture reasoning with implicit default knowledge. Consider the treatment of witness testimonies (the following discussion is adapted from Bex et al., 2007b). In a standard abductive approach, the relation between a witness testimony and its content must is represented as a causal rule, in which the testimony is regarded as caused by some other state or event. Usually, the testimony will be caused by the event to which it testifies (*p happened* → *witness w said "p"*). To be truly realistic this generalization should have some auxiliary conditions like "the *w* saw *p* happen" and "the *w* was interrogated" and so on, but for simplicity such conditions implicit will be left implicit. However, there may be other possible causes of the witness testimony. For instance, the witness could have hallucinated and now thinks *p* really happened (*w hallucinated that p* → *w said* "*p*") or the witness might want to protect the suspect and believes that testifying *p* will somehow help the suspect's case (*w wants to protect the suspect* → *w said* "*p*"). Given the above causal rules, there are three minimal hypotheses: {*p happened*}, {*w hallucinated that p*} and {*w wants to protect the suspect*}. Without further evidence, all three of these explanations follow as valid explanations for *w said p* and, as they all explain the one available explanandum, they are all equally good. However, intuitively it seems that in the absence of further evidence for the alternative explanations (that *w* hallucinated or that he wants to protect the suspect), they are not worth considering and that the usual explanation (that *p* really happened) should be accepted as the right one by default. In other words, the alternative explanations are exceptions to the general default statement that witnesses will usually speak the truth and they should therefore be assumed false as long as there is no evidence to the contrary. Thagard(2005) speaks in this connection of a "dual pathway model" of reasoning with testimonial evidence: he distinguishes a "default pathway" in which people almost automatically accept a testimony and a "reflective pathway" in which people build a causal model of the relevant knowledge and decide whether to believe the testimony by inference to the best explanation. People shift from the default to the reflective pathway when the content of the testimony is inconsistent with their current beliefs or when there is reason to doubt the credibility of the source. The problem with the standard approach to causal-abductive reasoning is that it forces the reasoner to always take the reflective pathway, since it forces to consider all alternative explanations, even if they are not supported by any further evidence.

Another drawback to modelling the evidential data as being caused by the events in the story is that it can be counterintuitive. In some cases, modelling knowledge as a cause – effect relation is very natural; for example, "hitting someone over the head with a hammer may cause brain damage" or "getting into a car with dirty shoes may cause shoeprints (in the car)". In other situations, however, the generalization expressing the relation is more contrived; for example "seeing someone get hit over the head with a hammer and being asked by a police officer whether one saw something causes one to testify that *I saw someone get hit over the head with a hammer*" or "if DNA profiles A and B match then an expert will testify that DNA profiles A and B match". Section 3.1.4 already referred to research (van den Braak et al., 2008) which shows that in the case of "testimonial knowledge", that is, information from testimonies and evidential documents, people find it significantly harder to interpret causal relations than they find it to interpret evidential relations.

An aspect of the story-based approach which is underdeveloped is exactly how stories should be rationally compared. Pennington and Hastie say that their principles determine the acceptability of a story in that they determine how the confidence in a story can be increased or decreased, but they do not say to what extent the various criteria influence this confidence and they do not precisely define the criteria. In the previous section, these criteria for the coherence of a story were more precisely defined but it is still unclear exactly how these criteria should be used in the dialectical process of testing multiple hypotheses. This is in contrast with the argument-based approach, for which several ways in which arguments can be attacked and defeated were discussed in Section 3.1.2, in which it was shown that there is a clear way of modelling the outcome of dialectical reasoning with arguments. Reasoning with stories or explanations has not been developed in this way and the criteria for comparing explanations proposed by more formal work in diagnosis and abductive inference to the best explanation are, as Thagard and Shelley (1997) have argued, often too simple for a complex domain such as reasoning with evidence.

Because stories and story schemes are, like generalizations, a way of talking about general commonsense knowledge of the world, there are also dangers of stories are comparable to the dangers of generalizations. Stories can, for example, sneak in irrelevant facts, focus attention on the actor rather than the act, appeal to hidden prejudices, tell the course of events in emotionally toned language and so on (Anderson et al., 2005 pp. 281–282). These dangers are partly obviated by the coherence requirements of the story-based approach, by ensuring that a story conforms to a plausible story scheme and simultaneously requiring that the causal connectivity and the consistency of the story is in order. The refinement of generalizations (Figs. 3.17 and 3.18) allows the reasoner to build a clear causal structure and thus improves the coherence of the story. However, a major shortcoming here is that it is impossible to reason *about* the causal generalizations or the story in general. In the argument-based approach, there were multiple ways of denying or questioning the use of unsafe generalizations, which is impossible in a purely story-based approach. In other words, there is no true cognitive consensus can about the story schemes and causal relations that have been (implicitly) assumed in the story. Wagenaar's et al. Anchored Narratives approach (Wagenaar et al., 1993) proposes a kind of dialectical process for determining the plausibility of the generalizations in a story, but it does not involve reasons for and against some piece of knowledge.

To summarize, stories provide a natural and cognitively plausible way of reasoning with several more complex hypotheses in a case. More holistic structures such as story schemes allow for quick construction of hypothetical scenarios and using the causal information and information about motives and actions contained in these scenarios they can be further developed and also be used to guide the search for evidence. However, the evidential data does at present not have a clear place and using only causal links to connect the data to the story explaining it is can be counterintuitive. The immediate dangers of stories are lessened by requiring that a story is coherent and plausible and hence choosing a hypothesis in a purely story-based approach can be considered a rational activity. In this sense I disagree with Wigmore in that stories cannot be used for the rational analysis of a case.[16] However, the possibilities for testing and comparing stories need to be further refined in order for the story-based approach to be properly rationally well-founded; it is at present impossible to reason about the generalizations or story schemes that underlie the story, and the dialectical process of inference to the best explanation has also not been fully developed.

[16] Wigmore argued that the (argumentative) chart method is the only "thorough and scientific method" for analysing evidence.

# Chapter 4
# A Hybrid Theory of Stories and Arguments

At the beginning of this book it was argued that we are looking for a theory that is both natural and also rationally well-founded. This theory should be flexible as to allow for the various kinds of reasoning in the contexts of discovery and pursuit whilst not getting too complex so that the results in each phase are relatively easy to understand. The theory should on the one hand allow for a careful testing of complex hypotheses and the evidence supporting these hypotheses but on the other it should provide a broader overview, both of what might have happened and of the available evidential data.

The argument-based approach allows for a thorough and rational analysis of the evidential data in a case. The generalizations used in the reasoning can be supported and attacked in multiple ways and the interaction between the various pieces of evidential data can be modelled in an attack graph. Due to its more atomistic nature, the argument-based approach is unsuitable for giving a clear overview of the various hypotheses about what happened in the case. Furthermore, not all aspects of causal reasoning such as the prediction of unexpected effects can be found in the purely evidential argument-based approach. Finally, while the idea of evidential arguments and generalizations is based on intuitive and natural concepts, the true usefulness of constructing and reasoning with such arguments in the process of proof has yet to be shown through empirical research.

In contrast, the story-based approach has been extensively empirically tested and is known to conform with the way humans reason about evidence and scenario's in a criminal case. A story allows for an easy overview and knowledge structures such as story schemes allow for a holistic way of reasoning with motives and actions. The interaction between the various elements of a story about what happened in a case can be explicitly modelled using causal relations. Because of the causal nature of stories, combining explanatory and predictive reasoning is easy and so called *storyconsequences* (i.e. consequences that might have causally followed from the story) can point to new evidence. However, in the story-based approach the evidential data does not have a clear place and therefore its credibility and relevance for the various elements of the hypotheses in a case cannot be checked as easily as with the argumentative approach. Furthermore, a purely story-based approach does not allow for reasoning *about* a story, that is, it does not allow for reasoning about its internal causal coherence.

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_4, © Springer Science+Business Media B.V. 2011

**Table 4.1** Two approaches to reasoning with evidence

| Arguments | Stories |
|---|---|
| Evidential | Causal |
| Atomistic | Holistic |
| Reasoning with and about evidential data | Reasoning about actions and events |
| Logically and conceptually well-developed account of the rational and dialectical process of proof | Relatively undeveloped, especially the dialectical aspects of comparing alternative stories |
| Only scantly empirically tested in the process of proof | Extensively empirically tested in the process of proof; shown to be a natural way of reasoning about criminal cases |

Table 4.1 briefly sums up the features of the two approaches. Note that here the argument- and story-based approach have been clearly separated and are presented as two wholly different approaches. It can be argued that this is sometimes an artificial distinction and that in practice, the distinction between stories and arguments may not always be clear. For example, when reasoning about a piece of evidence and its credibility are we comparing alternative causal explanations (i.e. very short stories) for what might have caused this evidence, as in Fig. 3.26, or are we attacking an evidential argument, as in Fig. 3.6? Similarly, when we are inferring a motive for an action, are we performing abductive practical reasoning with arguments or are we performing abductive reasoning with the story scheme for intentional action?

The way in which each approach is modelled also influences the advantages and disadvantages that can be attributed to each approach. For example, an argument-based approach in which only more abstract argument frameworks (as in Fig. 3.9) are used can be said to be more holistic, as it provides an overview of how all the evidence in the case interacts while sacrificing the details as to how exactly the events follow from the evidential data. Similarly, a story-based approach in which all the causal links between the events and the evidence are rendered in detail can be said to be more atomistic than holistic.

However, sometimes the distinction between an argument and a story is much clearer. The sequence of events in Fig. 3.16 is clearly not an argument and the inference from evidence to an event in Fig. 3.1 has, considered on its own, nothing to do with a story. Thus in an evidential argument-based approach it would not be possible to convey the information contained in Fig. 3.16 and a causal story-based approach would not allow us to evidentially infer one simple event $E$ from a piece of evidence $E^*$. In practice the distinction is similarly noticeable: for example, a judge considering a single event and the evidential data for and against this event is engaged in atomistic and more argumentative reasoning whilst an investigator trying to determine what course of events led to the current situation is comparing various stories in a more holistic way.

Bex and Verheij (2009) have argued that stories and arguments are in a sense "communicating vessels"; in some instances a causal, holistic and more story-based approach works best and in other instances an evidential, atomistic and argumentative approach is the most natural. Which of the approaches is used depends not only on the situation but also on the reasoner. Hence a *hybrid approach* that combines the arguments and stories provides the most expressive and flexible theory for legal evidential sense-making. This combination of the two approaches, in which the various hypotheses in a case are modelled as stories and where the evidential data is connected to the events in the story by evidential arguments, solves the problems of the individual story – and argument-based approaches. Stories can be used for the quick and creative construction of intelligible hypotheses about what happened in a case and at the same time arguments can be used to support these stories with evidence and to reason about the plausibility and coherence of the stories in greater detail. In this way the evidential data has a clear place and can be connected to the various hypotheses in an intuitive way. In the following section, the features of the combined theory will be described. Note that all the features of the individual story – or argument-based approaches remain intact in this combined approach; combining the two modes of reasoning only extends the possibilities.

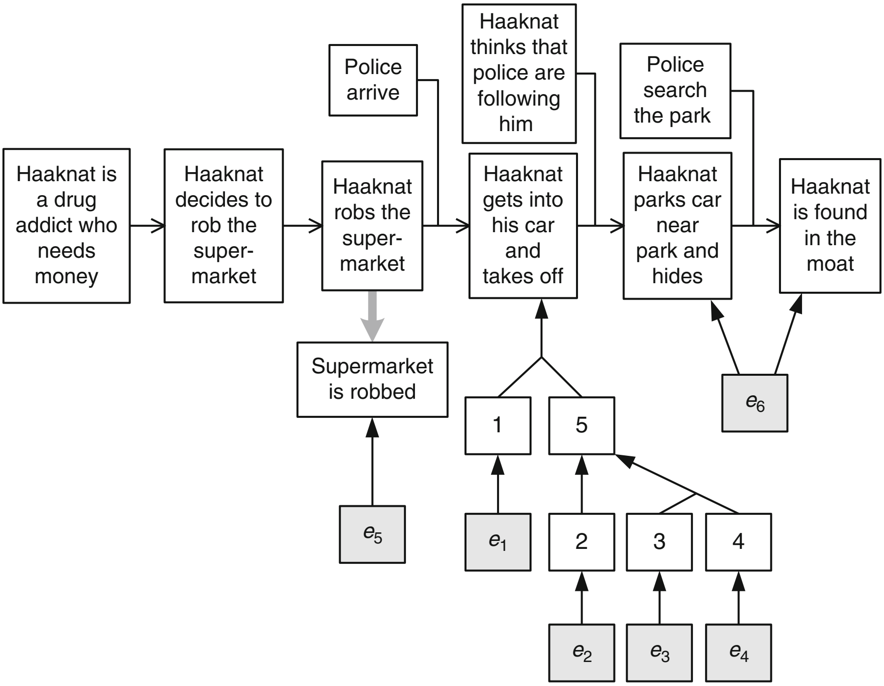
## 4.1 Combining Stories and Arguments

In the hybrid theory,[1] stories are modelled as simple causal networks. These stories causally explain the explananda in a case and hence can be regarded as the possible hypotheses about what happened. The evidential data are represented as separate propositions and from this data, the states and events in the stories can be inferred through evidential reasoning. The distinction between a piece of evidence $E^*$ and the event $E$ is thus preserved and it is possible to reason about the individual pieces of evidence by, for example, undercutting the inferences from $E^*$ to $E$.

The combination of two types of reasoning is best explained by giving an example. In the hybrid theory, the original Haaknat story (see p. 62 and Fig. 3.16) can be supported by the evidence in the way shown in Fig. 4.1. This story explains the explanandum that "Haaknat hides in a moat" and also the explanandum that the supermarket is robbed. In Fig. 4.1, all the pieces of evidence are now clearly represented and the argument from Fig. 3.4 (for the conclusion that Haaknat got into the car, see p. 40) now supports the story from Fig. 3.16 (that Haaknat robbed the supermarket)

In the purely story-based approach, the evidential data itself were considered to be the *observations* that had to be explained. In the hybrid approach, it is usually not the data itself but the observations that can be derived from this data that have to be explained. Take the distinction between an evidential datum $E^*$ and the event that can be inferred from the datum $E$ (see Section 2.1.2 on evidence). In the hybrid approach the observation that has to be explained will usually be the event $E$, whilst

[1] Previous versions of this hybrid theory were presented in Bex et al. (2006, 2007a, b) and Bex and Prakken (2008).




| | |
|---|---|
| $e_1$: Witness $W_1$ testifies that 'I saw someone looking like Haaknat get into the car'. | 1: Witness $W_1$ saw someone who looked like Haaknat get into the car. |
| $e_2$: Expert $E_1$ testifies that 'the DNA profile of the hair in the car matches Haaknat's DNA'. | 2: The DNA profile of the hair in the car matches Haaknat's DNA. |
| $e_3$: Police report: 'the shoeprints found in the car were of Runner shoes'. | 3: The shoeprints found in the car were of Runner shoes. |
| $e_4$: Police report: 'Haaknat wore Runner shoes when he was apprehended'. | 4: Haaknat wore Runner shoes when he was apprehended. |
| $e_5$: Witness $W_2$ testifies that 'I saw someone with a gun rob the owner'. | 5: Haaknat has been in the car at some time. |
| $e_6$: Police report: 'Haaknat was found in the moat and his car was nearby | 6: Haaknat was the man who got into the car. |

**Fig. 4.1** The Haaknat story and its associated evidence in the combined approach

in a purely story-based approach the observation will be the data $E^*$. Thus the arguments based on evidential data provide the observations that can be explained by the stories. As a consequence, in the hybrid theory the explananda should be supported by non-overruled arguments. In other words, only states or events about which there is no doubt that they happened should be explained by a particular story. If, for example, a witness reports a crime but it turns out that the witness for some reason does not speak the truth, there is no crime and hence the event to which he testified does not have to be explained.

While a requirement for an observation in the explananda is that it follows from a justified argument, this does not mean that all events that follow from justified arguments are automatically explananda that *have to be* explained. Recall that the choice of explananda is dependent not only on the evidence but also on the position of the person giving the explanation and on what is at the moment considered to be the best explanation (see the discussion on different explananda in Section 3.2.4). For example, Haaknat does not have to explain who robbed the supermarket if it is clear that he did not do it. Also, if there is overwhelming evidence that someone else than Haaknat robbed the supermarket it makes no sense to find a story which incorporates "Haaknat is a drug addict", even though Haaknat's addiction may be without question (i.e. follow from a justified argument). Thus the choice of explananda depends on the case as a whole and not just on what the justified conclusions are.

In the hybrid theory, the basic principles of the story-based approach and abductive inference to the best explanation still apply. That is, merely explaining the explananda with a coherent story that conforms to the evidence is not enough: alternative explanations should also be constructed and compared according to their coherence and the extent to which they conform to the evidence. Various criteria for comparing stories, such as the extent to which they conform to the evidence and their coherence, were discussed in Section 3.2.4. The arguments in the hybrid theory allow for more detailed and precise definitions of these criteria. For example, arguments can be used to support and attack events in a story. Furthermore, arguments can also support and attack causal generalizations in a story in the same way as arguments support or attack evidential generalizations (see Section 3.1.3). Basic principles of the argument-based approach also apply and can thus influence the comparison of stories. For example, evidential arguments that support a story can be defeated by other arguments, thus decreasing the evidential support of the story.

## 4.2 Evidential Support, Contradiction and Gaps

In the hybrid theory the evidential data is not directly causally explained by stories but rather the evidential data *supports* the stories: the conclusions of arguments based on the evidential data are states or events in a story. Because of this, the extent to which a story conforms to the evidence is called the total *evidential support* of a story.[2] Essentially, this evidential support is all pieces of evidence that support some event or causal relation in a story. For example, the story in Fig. 4.1 is supported by $E_1$ through $E_6$.

The evidential support of a story concerns not just the evidence that supports a state or event in the story: it is also possible for arguments based on evidence to support a causal generalization. For example, causes of death are often supported by pathologists' reports and the specific link between the cause (e.g. "the

[2] This support is similar to what Pennington and Hastie call *evidential coverage* (cf. Section 3.2.3 on explaining the evidence). The difference is that in the story-based approach, the story explains (i.e. *covers*) the evidential data itself whereas in the hybrid approach, the story explains the observations that are *supported* by evidence.

man was hit on the head") and its effect (e.g. "the man had a concussion") is often explicitly mentioned in such reports. In this way, a causal relation between two states or events is supported by evidence and this can be expressed by having the evidential argument support the causal generalization that expresses the relation.

Arguments can be attacked and defeated by other arguments they can be justified, defensible or overruled. Thus, counterarguments can influence the evidential support of stories: data that supports a story through an overruled argument does not count towards a story's total support. In such a case, the connection between the evidence and the story is severed because the evidence is in some way unreliable. For example, the argument from Fig. 3.6 (that the expert used obsolete methods to determine the DNA match) undercuts the argument from $e_2$ (the expert testimony) to the Haaknat story in Fig. 4.1 and thus the evidential support of the story is decreased.

The opposite of evidential support is *evidential contradiction*: the set of all pieces of evidence that contradict some element (i.e. a state, event or causal relation) in a story.[3] As with support, a piece of evidence can contradict a state or event in the story but also a causal generalization expressing one of the causal relations in the story. As with evidential support, overruled arguments are never considered strong enough to influence the extent to which a story conforms to the evidence. As an example of evidential contradiction, consider again the argument from Fig. 3.5 for the conclusion that Haaknat did not get into the car. This argument attacks the conclusion of the argument from Fig. 3.4 (that Haaknat got into the car) and thus it contradicts the event "Haaknat gets into his car and takes off just as the police arrive" in the story.

The above example shows that evidential support and contradiction are closely related. The argument for the conclusion that "Haaknat did not get into the car" (Fig. 3.5) contradicts the story and, supposing it is not overruled, increases the story's evidential contradiction. Now, this argument may also defeat outright any argument for the opposite conclusion; for example, the argument for the conclusion that "Haaknat got into the car" (Fig. 3.4). This argument that is defeated is also based on evidence and counts towards the evidential support of the story; defeating it then decreases the evidential support for the story. If the argument for "Haaknat did not get into the car" is justified then the argument based on $E_2$, $E_3$ and $E_4$ is overruled and hence the evidential support of the story in Fig. 4.1 is lessened at the same time as the evidential contradiction is increased.

Another important feature of stories that is related to evidential support or contradiction is that of *evidential gaps* in a story. According to Tillers (2005), such gaps are hypothesized events for which there is no direct evidence and which therefore have to be inferred from other circumstances for which there is evidence. One way of creating plausible circumstances for an event involves fitting the event in a coherent

[3] Similar to Pennington and Hastie's requirement that a story is consistent with the evidential data (cf. Section 3.2.3 on evidence in the story based approach).

story that is supported by enough evidential data, in other words: a good and true story. In this way, the circumstances detailed in the story make it more plausible that the events in question happened. This is the "gap-filling" function of stories that has been mentioned before, where gaps in the evidence are filled with events that fit the total picture painted by the story. In the case of the story in Fig. 4.1, there is only circumstantial evidence for the event that Haaknat committed the robbery, that is, there is no direct evidence (e.g. a testimony or CCTV images) from which it can be directly inferred that it was Haaknat who robbed the supermarket. This event can therefore be characterized as a gap-filler that fills the evidential gap of the identity of the robber.

So an evidential gap is a state or event for which there is no direct evidence. Events which are the conclusion of an overruled argument can be considered evidential gaps because there is no accepted argument based on evidence that says anything about these events. While the definition of a evidential gaps is thus dependent on the definition of evidential support, the notions are clearly different: in Fig. 4.1, the story has full evidential support (i.e. it is supported by all the evidence in the case at hand) while it still has a number of evidential gaps. The number of gaps in a story essentially represents the extent to which a story is supported by evidence: the more gaps constitute a story, the less it is based on evidential data. This is important to keep track of so that we do not run the risk of a good story (i.e. a coherent story with a large number of gaps) pushing out a true story (i.e. a less coherent story with a small number of gaps).

Note that evidential gaps concern only states and events that are not supported and that an unsupported causal generalization does not count towards the total number of evidential gaps. The reason for this is that one of the ideas behind reasoning with stories is that not all causal relations between events have to be explicitly mentioned, much less supported by evidential data. There are cases in which a certain unsupported causal relation is an important evidential gap; for example, in a tort case ideally evidence is presented from which it follows that the action of one party caused the damages to the other party. However, for the moment I will not consider such special cases. Remember that this does not mean that the causal connectedness of a story is not important and that it is still possible to reason about (evidence for) the causal relations in a story.

## 4.3 Story Coherence in the Hybrid Approach

Recall that there are three criteria for determining the coherence of a story (Section 3.2.4): the story has to conform to a plausible story scheme, it has to be internally plausible (i.e. its events and causal relations should be plausible) and the story should be internally consistent. Here, plausibility is a relative notion that depends on the current cognitive consensus about the stock of knowledge. In the hybrid approach this cognitive consensus can be reached by arguing about the stock of knowledge. Just as it is possible to reason about the plausibility of the generalizations in an argument, the plausibility of a particular story can be discussed by giving arguments for and against the plausibility of the underlying causal generalizations and by arguing that a story conforms to a plausible story scheme.

In this book, coherence and plausibility are notions that can be established *independently* of the evidential data: arguing about the coherence of a story is not done by arguing with evidence, but rather by arguing with commonsense knowledge that is part of our stock of knowledge about the world. The reason for this is that when there is evidence for a particular part of a story, its coherence is of secondary importance: even a highly implausible story can be believed if there is enough evidence for it. If, for example, an event is supported by evidential data, we do not need to reason with our commonsense knowledge in order to make this particular event or generalization more plausible. However, in cases where there is no evidential data supporting an event in a story it is important to consider whether the event is inherently plausible, that is, if it can be inferred from our stock of general commonsense knowledge. Similarly, if an event is already contradicted by evidential data, we do not need to show that the event is inherently implausible because it does not conform to our general knowledge

### *4.3.1 The Plausibility and Consistency of a Story*

An important aspect in the coherence of stories is a story's plausibility. That is, the states, events and causal relations that are not based on evidence should be plausible in that they follow from our stock of knowledge. In some cases, such events or relations are implicitly assumed. However, in the hybrid theory points of doubt in a story can be exposed in the same way as points of doubt in an argument can be exposed and thus the total plausibility of the story can be explicitly reasoned about. The *plausibility* of a story is then the extent to which the events and the generalizations expressing causal relations are supported by *explicit arguments* based on the stock of knowledge. So an event or causal relation which is explicitly supported by an argument from general knowledge is more plausible than an event or relation which is implicitly assumed. The rationale behind this is that if an explicit argument is given, the cognitive consensus about the assumed event or causal relation is more easily reached because explicit arguments can be tested in the dialectical process. In a similar way, the *implausibility* of a story is the extent to which the events and the generalizations expressing causal relations are attacked by explicit arguments based on the stock of knowledge. Whilst, for example, an implicitly assumed generalization expressing a causal relation can also be implausible, this is not evident unless an argument contradicting the relation has been given. Note that arguments based on evidence do not directly increase or decrease a story's plausibility, as they are already directly taken into account with evidential support and contradiction.

The plausibility of a story depends in part on the plausibility of its evidential gaps, that is, the intrinsic plausibility of events and causal relations in the story for which there is no evidence. Such events can be supported or contradicted by arguments from the stock of knowledge. For example, a story which assumes that there were US presidential elections in 2008 is supported (through an evidential inference) by the general knowledge that there were elections that year. On the other hand, an explanation that speaks of aliens robbing a supermarket is implausible because “aliens do not exists (on earth)” is most likely in our stock of knowledge and an evidential arguments based on this commonsense knowledge can attack the story.

Another part of the plausibility of a story is the plausibility of its (implicit) causal relations. In the examples of stories in this book, often a somewhat naïve or ad hoc interpretation of causality is used, where a causal link does not represent a much stronger relation than temporal precedence or some other supposed cue to causality. Take, for example, the story in Fig. 4.1. In this case Haaknat hides not only because he thinks the police are looking for him but also because he robbed the supermarket and hence it can be argued that there needs to be some kind of direct causal connection between the robbery and the fact that Haaknat hides in the moat. It can also be argued that some of the causal relations are somewhat far-fetched: for example, not all drug addicts who need money rob supermarkets. This interpretation of causality does not pose a problem in the hybrid theory, because the generalizations expressing the causal relations can be called into question using arguments. If, for example, some relation between two events in a story (e.g. temporal precedence or correlation) is falsely interpreted as a causal relation, this can be denied with an argument, thus increasing the story’s evidential contradiction (if the argument is based on evidence) or implausibility (if the argument is based on general knowledge). This approach to causality also allows for the causal structure of stories to initially remain relatively simple, so that they keep providing a good overview of the case; the causal structure of a story is only made more complex when there is reason to doubt this structure.

As an example, take the causal relation between “Haaknat is a drug addict who needs money” and “Haaknat decides to rob the supermarket”. This causal relation can be attacked and supported by commonsense knowledge in various ways (Fig. 4.2). Like in reasoning with evidential generalizations, with causal generalizations we can attack the inference by providing an exception to the

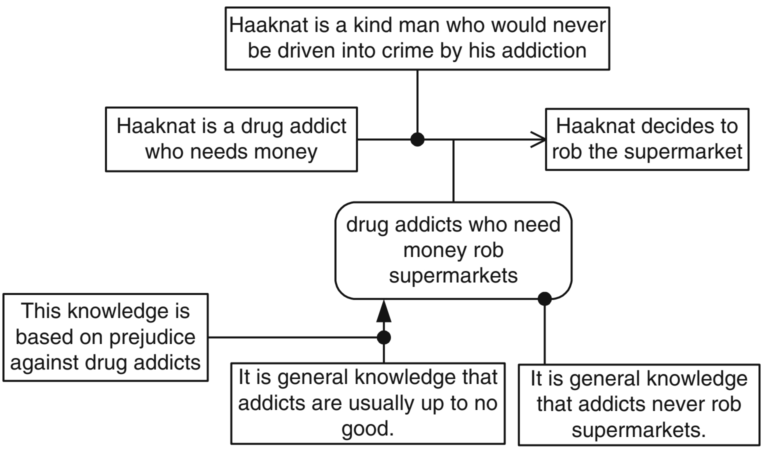



**Fig. 4.2** Supporting and attacking a causal generalization

particular instance of the generalization ("Haaknat would not do this") or we can attack the generalization itself ("drug addicts do not rob supermarkets"). It is also possible to attack the defeasible derivation from the source ("the general knowledge is based on prejudice"); here, the plausibility of the story is not directly attacked. Rather, the plausibility is decreased by overruling an argument that increases the plausibility. Notice that in Fig. 4.2 all arguments are based on information from the stock of knowledge. While it is perfectly possible to give an argument based on evidence for or against a causal relation (e.g. research which shows that supermarkets are often robbed by people who are drug addicts), discussions about the causal plausibility of a story are often based on commonsense knowledge and are about finding a cognitive consensus about the stock of knowledge. Furthermore, arguments based on evidence do not directly influence a story's plausibility, of which Fig. 4.2 is intended to be an example.

While arguments based on evidential data that support or attack a story do not *directly* count towards a story's causal plausibility, they can indirectly influence the causal plausibility by defeating an argument not based on evidence. Say that from the example in Fig. 4.2 the only argument that has been given is the direct argument against the generalization (that it is general knowledge that addicts do not rob supermarkets). If this argument is not attacked by another argument it is justified and hence decreases the causal plausibility of the story. Now, if an argument based on evidence (e.g. research which shows that most drug addicts are petty criminals) is given and defeats the argument against the generalization, the plausibility is increased because the argument appealing to general knowledge is now overruled.

A second criterion for story coherence is *consistency*, which concerns internal contradictions in the story. With this criterion, it is not the extent to which a story is consistent but rather the question *if* a story is consistent. In other words, a single inconsistency essentially causes a story to be incoherent. For example, if the Haaknat story in Fig. 4.1 were to contain an event that "Haaknat was found in a tree", then the story is inconsistent (as a person cannot be in a tree and a moat at the same time) and should not even be seriously considered as an explanation of the explananda. Contradictions are often implicit and determining consistency involves making inferences before it is shown that there is a contradiction. In some cases, these inferences are seemingly obvious (e.g. if someone is in a tree he cannot be in a moat). In other cases, they involve more complex defeasible causal prediction. For example, the event that "Haaknat robs the supermarket" does not only cause Haaknat to flee from the police, but it will usually also cause the money to be gone from the safe or the checkout registers. If the story contains the state that the money is still in the safe, the story may be considered inconsistent because an expected cause of one of the events in the story contradicts a state of affairs in the story.

### 4.3.2 The Completeness of a Story

In addition to the plausibility, a story's coherence also depends on its *completeness* (see Section 3.2.2 on story schemes and Section 3.2.4 on the criteria for determining the coherence of a story). That is, the story should complete a plausible story scheme by corresponding to all the elements of the scheme. In the hybrid theory, this completeness can be influenced in various ways. Firstly, it can be argued that a particular story scheme is implausible by attacking the scheme with an argument. Because a scheme is essentially a collection of connected causal relations, its relations can be attacked in the same way as the causal relations in a story can be attacked. For example, story scheme about a restaurant visit that contains the generalization "people who eat in restaurants usually leave without paying" can be attacked by an argument based on the commonsense knowledge that this is usually not the case. In addition, the correspondence relations between the story and the scheme should be expressed by plausible *abstraction generalizations* of the form "E *is a* SE", where E is an element in the story and SE is an element in the scheme. Using arguments, such abstraction generalizations can be attacked and thus the correspondence between the story and the scheme can be shown to be implausible.

In Fig. 3.23 it was already demonstrated that the Haaknat story completes the intentional action and the robbery scheme, both of which can be considered relatively plausible. Thus, the story conforms to the completeness requirement as originally defined by Pennington and Hastie. However, in previous work (Bex, 2009) I have shown that there are essentially two degrees in which a story is complete: a story *completes* a story scheme if for every element of the scheme it has a corresponding state or event. In other words, a story is complete if it "has all its parts". An additional requirement is that a story *fits* a story scheme; this is the case if all elements of the story correspond to some element in the scheme. Here a story is complete if it has no "loose ends" that do not fit the scheme. If, for example, an event "Haaknat was ordered by a higher power to commit the robbery" is added to the Haaknat story, the story still completes the robbery scheme. However, it does not completely match the robbery scheme as this scheme does not contain an element "*x* was ordered by higher power *z* to rob *y*". One way to have the story completely match the scheme is to extend the scheme with this element. However, this makes the new robbery scheme implausible: people are usually not ordered by higher powers to rob supermarkets.

## 4.4 Assessing and Comparing Stories

In the hybrid theory, there now are essentially six criteria for determining how good a story is. The extent to which a story conforms to the evidential data can now be broken down into evidential support, evidential contradiction and evidential gaps; the coherence of the story is determined by looking at the plausibility, the completeness and the consistency of the story. The combination of arguments and stories modelled as causal explanations allow these criteria to be defined in a specific way.

The criteria basically provide *critical questions* for the analysis of a story. The criteria phrased as questions provide typical sources of doubt for a combination of evidential data and a story, in the same way that critical questions associated with argumentation schemes can be seen as sources of doubt for a single, one-step defeasible inference. Three of these questions concern the extent to which the story conforms to the evidence:

- How much and which of the available evidential data supports the story? (*Evidential support*)
- How much and which of the available evidential data contradicts the story? (*Evidential contradiction*)
- How many and which events in the explanation are unsupported by evidential data? (*Evidential gaps*)

The other three questions concern a story's coherence and can essentially be answered without taking into account the evidential data in the case:

- How inherently plausible are the events and generalizations expressing the underlying causal relations in the story? (*Plausibility*)
- Does the story complete a plausible story scheme? Does the story fit a plausible story scheme? (*Completeness*)
- Are there elements of the story that contradict each other? (*Consistency*)

These criteria are essentially a more detailed and thorough description of Pennington and Hastie's criteria (Section 3.2.4) and the general criteria that apply to general inference to the best explanation (Section 2.3.1).

Given the above critical questions, a story can now be constructed and analysed through a combination of explanatory and predictive causal reasoning and evidential reasoning. Take as an example the Haaknat case; one of the explananda in this case is that Haaknat is found in the moat. The prosecution tells a story that Haaknat jumped in the moat because he was trying to hide from the police after he robbed the supermarket. This story explains why Haaknat was in the moat. The story's conformance to the evidential data is shown in Fig. 4.1; at the moment, the data $e_1$ through $e_5$ support the story. Important evidential gaps are the exact identity of the person who robbed the supermarket and the identity of the person who parked the car: in the story it is argued that Haaknat did both these actions but no argument is supplied from which it directly follows that it was actually Haaknat. The story seems not to be directly contradicted; however, assume that we have a testimony by Haaknat which says that he was not near the supermarket but rather in an argument with Bennie (Fig. 4.3).

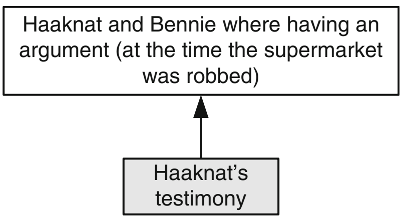



**Fig. 4.3** Haaknat's testimony about the argument with Bennie

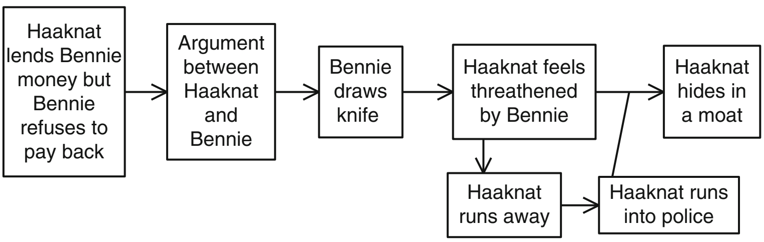



**Fig. 4.4** Haaknat's reason for hiding in the moat

This argument contradicts the story that Haaknat robbed the supermarket (Fig. 4.1): Haaknat could not have robbed the supermarket and have an argument with Bennie at the same time. The fact that Haaknat was doing something else can also be modelled as an alternative explanation for the explanandum that Haaknat was found in the moat (see Fig. 4.4). This shows that information about, for example, an alibi can be treated both as contradictory evidence and as an alternative explanation and is therefore an example of the fact that stories and arguments are essentially communicating vessels.

The story in Fig. 4.1 seems to be internally consistent. Reasoning about the story's internal (causal) plausibility is shown in Fig. 4.2. Note that when such arguments for or against causal generalizations are not given, the generalizations are simply assumed. In Fig. 3.23 it was shown that the Haaknat story matches the intentional action scheme and thus also the robbery scheme. Thus, if we believe these schemes to be plausible, the story is complete.

Once a basic story has been constructed, further explanatory and also predictive reasoning can be applied in order to further improve or worsen the story. For example, predicting story-consequences can worsen the story because new evidential gaps are created: if we predict that "Haaknat gets into the car and takes off" has as an effect that there must be fingerprints on the steering wheel and no such fingerprints are found, the total story becomes less supported by evidential data. However, if the fingerprints are found predicting that there are fingerprints may also improve the story because more of the evidential data supports it.

### *4.4.1 Comparing Stories*

The above criteria for assessing the quality of a story only concern a single story and its relation with the evidence. In order to compare explanations, one should also take into account alternative explanations. Josephson and Josephson (1994) accordingly define additional criteria that can also be phrased as questions:

- How many other stories are there that explain the explananda?
- How decisively does the current story surpass these alternative stories?
- How thorough has the search for alternative stories been?

In the Haaknat example, alternative explanations for the explanandum that Haaknat was hiding in the moat can be found. Above, it was argued that contradictory evidence can point to alternative explanations: if this evidence incompatible with one explanation, there must be another explanation with which it is compatible. Above, the prosecution's story in the Haaknat case was contradicted by Haaknat's own testimony that he was having an argument with Bennie at the time of the robbery. This testimony points to the alternative story in which Haaknat ran away because he and Bennie got into a fight:

This explanation can only supported by $e_6$ (the police report that Haaknat was found in the moat) and Haaknat's testimony that he ran away from Bennie. Evidence $e_1$ through $e_4$ *indirectly contradicts* the story because they support an alternative explanation. The conclusion that Haaknat got into the car could be extended to something like "Haaknat could not have been in a fight with Bennie because at that time he was getting into a car" and then $e_1$ through $e_4$ would directly contradict the story.

It can be argued that Haaknat's own testimony does not provide sufficient support for the story, as Haaknat clearly has a valid reason to lie about the events. He might have made up the story about Bennie because the police regard him as a suspect and he wanted to provide himself with a plausible alibi. If the arguments from Haaknat's testimony to the events in his story are hence undercut (because Haaknat's veracity is questionable), Haaknat's story is largely unsupported. This shows how argumentative reasoning influences the quality of a story: evidential data that is connected to a story through an overruled argument does not count towards the evidential support. Furthermore, it also shows that a single piece of evidence does not provide broad support for a story: if Haaknat's testimony is judged as not credible, most the story suddenly is unsupported.

As for the coherence of Haaknat's story, it seems that the story is not internally inconsistent. An important question is whether the story is internally plausible: because no additional evidence for the events Haaknat testified to is given, we should pay additional attention to the inherent plausibility of his story. Most of the assumed events and generalizations are reasonably plausible: arguments often start about money and such arguments can cause people to threaten others. However, the causal link "Haaknat feels threatened by Bennie and Haaknat runs into police $\rightarrow$ Haaknat hides in a moat" is dubious. Expressed as a generalization, this would read "Someone who feels threatened by another person and encounters the police will hide". This seems a strange generalization, because one would expect the person who was threatened to seek help from the police. So the story is less coherent because one of its internal causal generalizations is implausible. It can be argued that the fact that Haaknat feels threatened is not the cause of him hiding in the moat. However, as was shown in Fig. 3.28, the story is in that case not complete as then there is no clear psychological state that motivates Haaknat's behaviour.

An important question here is when one story is better than another story. One way to compare stories is to determine the extent to which the stories conform to the evidence with the first three criteria and the extent to which a story is coherent with the second three criteria and then provide an ordering on stories (i.e. *the more* *evidential data that supports the story, the better the story* or *the less evidential gaps the better the story*). This comparison could be done by simply counting, for example, the total number of evidential data that support story $S_1$ and $S_2$; if $S_1$ is supported by more data it is better than $S_2$ and vice versa. This provides an arguably crude but effective method of comparing stories: in the case of the Haaknat example, the prosecutor's story is clearly better supported by evidential data. Another way to compare stories would be to compare two stories using set inclusion: if the set denoting the evidential support of $S_2$ is a subset of the set denoting the evidential support of $S_1$, then $S_1$ is better. The idea behind this is that story $S_1$ only wins over story $S_2$ if $S_1$ is supported by at least the same evidence as $S_2$ and then some more. This method actually ensures that there is no explanation which is better than the one selected. However, a disadvantage is that in this way stories are often incomparable. In a real case, we are dealing with two or more stories that are supported by their own evidence. These stories express different views of what happened and are therefore rarely supported by the same evidence. As can be seen in the Haaknat example, the set of data supporting the prosecution's story $S_1$ is incompatible (i.e. it is not a sub – or superset) with the set of data supporting Haaknat's story. This is because when we have two alternative stories $S_1$ and $S_2$, $S_1$ will often be supported by evidence that contradicts $S_2$ and vice versa (as contradictory evidence often points to an alternative explanation, see above).

If the criteria are used to provide an ordering on stories, they should also be ranked according to their importance. On this subject, it should be noted that the extent to which a story conforms to the evidence should always be ranked as being more important than the coherence of the story. Recall from Section 3.2.5 on the evaluation of stories that the danger of a "good" story pushing out a "true" story is real. If a more supported story is, according to the ordering, always better than another story no matter the coherence of both stories, this danger is obviated.

Interpreting the various criteria in the more discrete, almost mathematical way as discussed here is not without problems. It is, for example, not always the case that if story A is supported by more evidential data than story B, story A is better than story B. If story A is supported by only one piece of evidence that is deemed highly credible and relevant and story B is supported by multiple pieces of evidence whose relevance to the main explananda is slight, we would not say that story B is automatically better because it has a higher evidential support. Furthermore, aside from the fact that a story's conformance to the evidence is more important than a story's coherence, the ranking of the various criteria is not trivial. What if, for example, one story is better supported than another story but also more contradicted?

The same types of problems encountered here were essentially discussed with regards to the strength of arguments in Section 3.1.2. The point is that reasoning with arguments and stories is always context-dependent and dependent on the knowledge available at the time the decision (for one argument or story) is made. Concepts such as "plausibility" are hard to quantify and simply "counting" the strength or degree of belief of an argument or story disregards this fact. For example, if we were to define plausibility as the number of causal generalizations in a story that are attacked by arguments, the plausibility of the Haaknat story is just as plausible as the plausibility of the prosecution's story. However, most people would agree that the prosecution's story is more plausible.

Thus the choice between stories is always made with regards to the content of the evidential data and the story. This is essentially the same as for comparing arguments: even though it can be precisely defined which arguments attack which other arguments, the ultimate choice about whether one argument defeats the other involves a judgement about the preference between two arguments. The reasons for this preference are often impossible to express in mathematical terms and can hence only be expressed in more context-dependent and substantive terms.[4]

A different way of using the criteria is by not interpreting them as providing hard-and-fast rules for the comparison of stories, but rather to regard them as providing guidelines for reaching a rational and well-thought-out decision about the facts of the case. Bex and Verheij (2009) give a list of *pitfalls* which are similar to the negative answers of the above critical questions (i.e. *the story is insufficiently supported by evidence*). They argue that the objective of the reasoner should be to avoid these pitfalls but that there is no absolute and objective measure that tells us when a pitfall has been avoided. In this way, these pitfalls serve as guidelines and not as hard requirements for a story like Wagenaar, van Koppen and Crombag's *universal rules of evidence* (Wagenaar et al., 1993, p. 231).

### *4.4.2 A Game for Inquiry Dialogue*

Recall from Chapter 1 that this book assumes a procedural notion of rationality: a belief or decision is rational if it is in agreement with the knowledge that has actually been considered (or should have been considered) in a proper procedure. A good way of testing a story is to see if it and its supporting evidence can stand against criticism in a dialectical inquiry. Recall that in argumentation such a dialectical process can be regulated and that and the structure and rules of a proper dialectical dialogue can be given as a protocol in a *dialogue game*. Dialogue games formulate principles for coherent dialogue between two or more players, and this coherence depends on the goal of a dialogue. In the argument-based persuasion game shown at the end of Section 3.1.2, the proponent made a claim which he had to defend, while the opponent's goal was to dispute this claim. The goal of the persuasive dialogue game was to resolve this difference of opinion in a fair and effective way. In a negotiation game (see Rahwan et al., 2004), the goal is not to get the other player to agree to some substantive claim but rather to negotiate about some "good". The

[4] On p. 41 it was shown that one way of solving the question which argument should be preferred is to allow for reasoning *about* the preferences between arguments or rules. In the same way, we could allow for such "meta-level" reasoning about the preferences for different stories (e.g. "my story is better than yours because it is supported by an important witness statement"). However, because of the complexity of the combination of stories and arguments, this direction is here not further pursued.

players, for example, can offer and request these goods and the game ends if one of the player's offers or requests has been accepted or the players have reached a compromise. The rules of a game ensure, for example, that all the moves that are made are relevant to the issue at hand. For example, in an argumentative persuasion dialogue the opponent may only directly attack claims made by the opponent and not start discussing something totally different.

In the process of proof, or at least in the process of proof as relevant for crime investigation, the players of the dialogue game have identical roles since they all want to find the most plausible and evidentially well-supported explanation for the explananda. A dialogue in the process of proof can be characterized as an *inquiry dialogue* (Walton, 1998). According to Walton, the overarching aim of an inquiry dialogue is to increase our knowledge; the players of the game collectively gather, organize and assess hypothetical stories and evidence. Walton (1998) further identifies two functions of an inquiry dialogue. The first is to *explain* why or how something happened and the second is to *argue* that one account of what happened should be accepted given the current evidence and other knowledge. In this way, an inquiry dialogue can be divided into three phases, which roughly correspond to the discovery, pursuit and justification phases from the process of proof. In the first phase, which corresponds to Walton's idea of explanation, the players each build their own explanation and concentrate on supporting and extending this explanation. In the second phase, which corresponds to Walton's *argumentation*, the players start critically analysing explanations, concentrating on attacking the other players' explanations while at the same time defending their own explanations by, for example, undercutting arguments that contradict their explanation. In the third phase, the players try to find a compromise.

Notice that because of the combination between explanation and argumentation an inquiry dialogue incorporates both a *cooperative* and an *adversarial* side. Whilst none of the players really wants to "win" (since the goal of the dialogue is to increase the knowledge about the case and the players all want to find the best explanation for the explananda), disregarding the adversarial part of the dialogue increases the risk of tunnel-vision. If no alternative has to be proposed and the explanation does not have to be called into question, the players' natural confirmation bias will lead them to focus on one explanation and disregard alternatives. The adversarial setting can be partly enforced by demanding that the players constantly aim to have an explanation which is better than the other players' explanations. In order to achieve this, they will then extend and support their own explanation and attack the other players' explanations.

Summarizing, a dialogue game that provides the rules for a typical inquiry dialogue in the context of reasoning with criminal evidence will ensure that the discussion about the various hypothetical stories and arguments based on evidential data is conducted in a proper and rational way. This means that the utterances by the players should at least be *relevant*. Furthermore, through the adversarial setting the players should be encouraged to improve their own explanations or disprove the other players' explanations, thus hopefully avoiding "tunnel-vision". The idea is that if a discussion in the process of proof is regulated by the rules of such a game, pitfalls will be avoided. The precise dialogue game will be fully defined in Section 5.5.

## 4.5 Evaluation

In this chapter, it was argued that the best way of reasoning in the process of proof is to combine stories and arguments into a *hybrid theory*. In the hybrid theory, stories can be used to causally explain the explananda. These hypothetical stories can be supported and contradicted by evidential arguments based on evidence or general world knowledge.

The hybrid theory combines the advantages of the separate story – and argument-based approaches. Argumentative discussions about individual elements of the case (i.e. pieces of evidence, generalizations, elements of stories) are possible. In this way, the evidence and its credibility can be thoroughly analysed and a cognitive consensus about the general knowledge underlying the reasoning from evidence to the story can be reached. Stories organize the evidential data and give an overview of the case. Furthermore, stories and story schemes can be used to abduce relatively complex hypothetical scenarios. Finally, the causal nature of stories allows for the prediction of story-consequences, new events for which evidence can be sought.

In addition to combining the advantages of the separate approaches, the hybrid theory also solves most of the problems of the story – and argument-based approaches. For example, it is now possible to reason *about* stories using arguments and thus ensure the stories coherence in a dialectical way. The assumed events and relations in a story can be debated using arguments and thus a cognitive consensus not only about evidential generalizations but also on the general knowledge underlying a story can be reached. Furthermore, the addition of evidential arguments to the hybrid theory allows for precise definitions of the various criteria that can be used to compare stories. Evidential support and evidential contradiction provide a clear notion of how a story can conform to the evidence and the plausibility and completeness of a story ensure that the story conforms to general commonsense knowledge about the world.

The criteria for determining the quality of a story can best be phrased as *critical questions* which can be asked for a particular hypothetical story during the process of proof (e.g. "How much and which of the available evidence supports the current hypothesis?"). Given the current procedural conception of rationality, asking these questions ensures that the process of proof is a rational process and we may expect that pitfalls such as confirmation bias or unclear justifications of hypotheses are avoided.

# Chapter 5
# A Formal Logical Hybrid Theory of Argumentation and Explanation

In Chapter 4, a hybrid approach to reasoning with criminal evidence was proposed. In this chapter several notions from formal logics were discussed (i.e. the dialectical status of arguments and causal-abductive reasoning). This chapter aims to concretize the hybrid theory by defining a formal logical version of the hybrid theory.[1]

The basic idea of the formalized hybrid approach is as follows. A logical model of abductive *inference to the best explanation* (IBE) takes as input a causal theory (a set of causal rules or generalizations) and a set of observations that has to be explained, the explananda, and produces as output a set of hypotheses that explain the explananda in terms of the causal theory. The combination of hypotheses and causal theory can be seen as a story about what might have happened. These hypothetical stories or *explanations* can then be compared according to their coherence and the extent to which they conform to the evidential data in a case. This data is connected to the stories by defeasible arguments which can be attacked and defeated. Defeasible arguments are also used to attack explanations: the causal rules of the theory are not just given but their validity and applicability can become the subject of an argumentation process. In this way the explanations can be supported or attacked by the arguments.

This chapter is structured as follows. In Section 5.1 the features of the logic underlying the hybrid theory will be presented. Section 5.2 discusses the argumentative part of the hybrid theory and shows how arguments based on evidence can be constructed and attacked. Section 5.3 gives a brief and general overview of causal model-based reasoning and shows how the basics of the causal part of the hybrid theory can be modelled given the current defeasible logic. Section 5.4 discusses the hybrid theory, that is, the combination of the argumentative and the causal parts from the preceding sections. In Section 5.5 the formal dialogue game for constructing and testing stories and evidence is proposed. Finally, in Section 5.6 an extended example is given in which the formal hybrid theory, its semi-formal variant and the dialogue game are illustrated.

[1] The formal theory has been published in a more condensed form as (Bex et al., 2010).

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_5, 

## 5.1 A Defeasible Logic

Logic basically consists of a combination of a formal object language and a notion of valid consequence expressed in a metalanguage. The object language, from here on simply referred to as the logical language, usually is the standard language of first-order logic or an extension of it. The metalanguage, which allows us to reason about the formulas in the object language, contains the rules of inference that determine when and how formulas in the object language may be inferred from each other. An example of such an inference rule is the well-known rule for modus ponens, which allows us to infer the object-level formula $\psi$ from the formulas $\varphi$ and $\varphi \rightarrow \psi$ (where $\rightarrow$ is the standard logical material implication).

Before continuing the discussion of the current logic, first some remarks about logical rules and generalizations have to be made. Very broadly, a rule is considered to express the relation between a conclusion and the proposition that is considered to be the reason for that conclusion.[2] Note that here a distinction should be made between a rule at the object-level like "if $\varphi$ then $\psi$" (formally represented as $\varphi \rightarrow \psi$) and a metalinguistic rule of inference such as "$\varphi$. If $\varphi$ then $\psi$. Therefore $\psi$' (modus ponens, formally represented as $\varphi, \varphi \rightarrow \psi \vdash \psi$). This example shows that some rules of inference make use of a rule-like conditional in order to infer a conclusion. Another example of such an inference rule in classical logic is modus tollens (inferring $\neg\varphi$ from $\varphi \rightarrow \psi$ and $\neg\psi$). Not all inference rules, however, have a rule-like conditional statement as one of their premises; examples of such inference rules from classical logic are $\wedge$-Introduction (inferring $\varphi \wedge \psi$ from $\varphi$, $\psi$) and $\vee$-Introduction[3] (inferring $\varphi \vee \psi$ from $\varphi$). The similarity between object-level rules and meta-level rules becomes apparent if we consider that this last inference rule can in a sense be phrased in natural language as "if $\varphi$ then $\varphi \vee \psi$". Similarly, modus ponens can also be rephrased as "IF ($\varphi$, if $\varphi$ then $\psi$) THEN $\psi$".

Consequently, a conditional generalization of the form "if…then…" can be modelled as an object-level conditional or as a metalinguistic inference rule. For example, in Prakken's argumentation system (Prakken, 1997) the generalization "if something is a bird then it can (presumably) fly" would be represented as an object-level rule $r$: $\mathsf{Bird}(x) \Rightarrow \mathsf{Canfly}(x)$, where $\Rightarrow$ is defeasible implication, a variant of the material implication. If we then also accept that Bird(Tweety), we can infer that Tweety can fly by applying a defeasible variant of modus ponens to these two formulas. In Reiter's default logic (Reiter, 1980), however, the generalization would be modelled as a (domain-specific) inference rule called a default. This default $d$: $\mathsf{Bird}(x)\!: \mathsf{Canfly}(x)/\mathsf{Canfly}(x)$, informally stands for "if $\mathsf{Bird}(x)$ holds and $\mathsf{Canfly}(x)$ may be consistently assumed, then $\mathsf{Canfly}(x)$ may be inferred". Here, accepting Bird(Tweety) also allows for the inference of Canfly(Tweety) using the default $d$.

---

[2]This conception of *rule* is based on, but not the same, as Hage and Verheij's definition of a rule, see (Verheij et al., 1998).

[3]Note that these rules may have different names in the various deductive systems. See (Gabbay et al., 1993) for an overview.

The above example also shows that inference rules can range from abstract to specific. Modus ponens is an example of an abstract inference rule. Another example of a abstract rule is the defeasible variant of modus ponens : "$\varphi$. If $\varphi$ then $\psi$. There is no exception to the rule *if* $\varphi$ *then* $\psi$. Therefore $\psi$". This rendition of the rule is taken from Verheij (2003b), who calls it *modus non excipiens*; Walton and Reed (2002) call Verheij's rule defeasible modus ponens. Prakken (1997) and Prakken and Sartor (1997) define a similar rule. In his logic, Pollock uses a mix of abstract and specific inference rules, which he calls *reasons* (Pollock, 1987; 1995, Chapter 2). An example of an abstract rule is the statistical syllogism ($x$ is a $\varphi$ and $n$ percent of $\varphi'$s are $\psi'$s is a reason to believe that there is an $n$ percent chance that $x$ is a $\psi$). Other reasons are based on less abstract but still generally valid epistemic principles, such as perception ("having a percept with content $\varphi$ is a prima facie reason to believe $\varphi$") or memory ("Recalling $\varphi$ is a reason to believe $\varphi$"). Based on Pollock's logic, Bex and colleagues (2003) have defined reasons which can be viewed as domain-specific generalizations (e.g. "Witness $w$ says '$\varphi$' is a reason for believing $\varphi$"). Reiter's default about birds which was mentioned above is also an example of a domain-specific inference rule.

In sum, reasoning patterns range from abstract (e.g. defeasible modus ponens) to concrete and domain specific (e.g. the witness testimony reason). Furthermore, conditional generalizations of the form "if. . .then. . .", which are used in the evidential arguments, can be modelled in two ways: as conditional premises for an abstract (defeasible) modus ponens inference rule or as inference rules themselves. In the first case, a generalization is part of the logical object language (rule $r$ in Prakken's logic) and in the second case the generalization is a metalinguistic inference rule (default $d$ in Reiter's logic). Verheij (2003b) has argued that a logic of which the rules of inference are based on domain-specific generalizations can be seen as a *contextual logic* (e.g. Reiter, 1980; Pollock, 1995; Bex et al., 2003), while a logic which only has abstract reasoning patterns such as (defeasible) modus ponens as rules of inference can be seen as an *abstract logic* (e.g. Prakken and Sartor, 1997; Verheij, 2003a). A purely abstract logic is in a way more flexible in that the reasoner determines which generalizations or reasoning patterns he wants to use and is not constrained by the specific inference rules of a highly contextual logic. A contextual logic is less flexible but it better captures a specific context and the types of inferences made in the context and thus ensures a certain context-dependent standard of rationality by providing the reasoner with certain principles to guide his reasoning. The main practical difference between modelling a generalization as a conditional premise and modelling it as an inference rule is that the validity of an inference rule cannot be questioned or denied, whilst the validity of a generalization as a conditional premise can be questioned.

It is possible to have a combination of abstract and contextual inference rules, as there is a continuum ranging from truly abstract logics to completely contextual logics. In such a "mixed" logic, abstract rules of inference like defeasible modus ponens allow for flexible reasoning with all kinds of rules and generalizations and specific inference rules provide handles or guidelines the reasoner can follow. The current context of reasoning with evidence in criminal cases lends itself well to such a mixed approach. Oft-used generalizations the general validity of which cannot sensibly be questioned can be modelled as domain-specific inference rules and generalizations of which the general validity is questionable can be represented as conditional formulas in the object language. Exactly which generalizations can be considered safe enough to model as inference rules is something that will be discussed later in Section 5.2.1 and 5.3.1. First, the logical object language has to be formally defined.

The above discussion shows that, in addition to the connectives of standard first-order logic, a connective is needed to express conditional generalizations. Furthermore, the evidential data that is used as the input for our arguments should be labelled accordingly.

**Definition 5.1.1 (Language)**

*Let $\mathcal{L}$ be any first-order language.*

- *A* defeasible generalization *is an expression of the form*

  $g_i(\vec{t})$: $\varphi \Rightarrow_{\mathsf{T}} \psi$

  *where $g_i(\vec{t})$ is the generalization's name according to some convention and $(\vec{t})$ is a tuple that denotes the simple terms (*i.e. *variables and constants) in the generalization. $\varphi$ and $\psi$ are conjunctions of literals from $\mathcal{L}$ that denote respectively the* antecedent *and the* consequent *of the generalization. The subscript $\mathsf{T}$ is one of* $\{\mathsf{E}, \mathsf{C}, \mathsf{A}\}$ and *denotes the type of generalization, which is evidential* ($\mathsf{E}$)*, causal* ($\mathsf{C}$) *or abstraction* ($\mathsf{A}$)*. A generalization with open formulas (*i.e. *one that contains variables) is a scheme standing for all its ground instances.*
- *A piece of* evidential data *is of the form $e_i$: $\varphi$, where $e_i$ is the data's name and $\varphi$ is a wff in $\mathcal{L}$.*
- *The current* defeasible language $\mathcal{L}_{\mathrm{d}}$ *based on $\mathcal{L}$ is $\mathcal{L}$ extended with the set of all defeasible generalizations, the set of evidential data and the set of defeasible input data.*

The above definition is based on Prakken's (1997, p. 153) definition: Prakken, speaks of "defeasible rules" instead of generalizations and calls the combination of a first-order language and defeasible rules a "defeasible extension". In the rest of this book, logical formulas will often be paraphrased in a semi-formal way, so instead of the formula witness($w$) $\wedge$ testifies($w$, $x$) I will write witness $w$ testifies "$x$". Variables will be represented as letters $w$, $x$, $y$, $z$ and constants with letters $a$, $b$, $p$, $q$. Greek letters $\varphi$, $\psi$ and $\chi$ denote well-formed formulas (*wffs*) in the language $\mathcal{L}_{\mathrm{d}}$. In a generalization's name, $\vec{t}$ denotes any tuple of simple terms and denotes a tuple containing only constants. Furthermore, a generalization $g_i(\vec{t})$: $\varphi \Rightarrow_{\mathsf{T}} \psi$ can be referred to with its name $g_i(\vec{t})$ and, when the specific terms in a generalization are not important, also with $g_i$.

Now the defeasible implication $\Rightarrow$ can be used to express both conditional and non-conditional generalizations. The difference between evidential, causal and abstraction generalizations is denoted with a subscript: an evidential generalization is represented as $\Rightarrow_{\mathsf{E}}$, a causal generalization is represented as $\Rightarrow_{\mathsf{C}}$ and an abstraction generalization as $\Rightarrow_A$. As examples of a conditional evidential generalization, consider the following generalization from the argument in Fig. 3.3 on p. 38 (the chain of inferences leading from witness $W_1$'s testimony to the conclusion that Haaknat got into the car):

$g_L$(*w, x*): witness *w* saw someone who looked like *x* $\Rightarrow_E$ *w* saw *x*

Causal and abstraction generalizations can be expressed similarly; consider a causal generalization expressing a link in the Haaknat story in Fig. 4.1 (p. 86) and an abstraction generalization expressing a correspondence link in Fig. 3.22 on p. 68 (the correspondence between the intentional action and the robbery schemes):

$g_c$(Haaknat, supermarket): Haaknat decides to rob supermarket $\Rightarrow_C$ Haaknat robs supermarket

$g_a$(*x, y*): *x* robs *y* $\Rightarrow_A$ action

As can be seen, the logical language is sufficiently expressive to express generalizations of various degrees of specificity. The generalization "if a witness saw someone who looked like Haaknat, then the witness saw Haaknat" can be expressed as follows, where H is short for Haaknat:

$g_L$(*w,* H): witness *w* saw someone who looked like Haaknat $\Rightarrow_E$ witness *w* saw Haaknat

This generalization is essentially a more specific version of generalization $g_L$(w, x). With this more specific version it is only possible to reason about witnesses who saw someone who looked like Haaknat and not about witnesses who saw someone who looked like someone else than Haaknat. So the language allows for the expression of (case-)specific generalizations. Following Prakken (1997), assumptions which are not conditionals can also be expressed in the language as a generalization of the form $\Rightarrow_E \varphi$, which is shorthand for $\top \Rightarrow_E \varphi$:

$g_{dap}$: $\Rightarrow_E$ the belief that drug addicts often want to rob supermarkets is based on prejudice

$g_{gkL}$: $\Rightarrow_E$ it is general knowledge that "if a witness saw someone who looked like person *x* then the witness saw *x*"

The second general knowledge generalization has already been shown in Fig. 3.11 on p. 50.

Now that the logical language has been defined, the various inference rules that allow reasoning with this language need to be defined. In previous work (Bex et al., 2003, 2007a), the inference rules of classical logic were augmented with defeasible inference rules. This earlier work is based on Pollock (1995), who argues that reasoning can be done with two kinds of inference rules which he calls *prima facie reasons* and *conclusive reasons*, respectively. Conclusive reasons are the deductive inference rules from classical logic; prima facie rules, by contrast, are defeasible inference rules in that they only create a presumption in favour of their conclusion. Such a prima facie reason can be undercut: for each prima facie reason, Pollock explicitly defines its undercutters.

The question now is what kinds of prima facie reasons should be included in the current contextual logic for reasoning with evidence. In order to reason with defeasible generalizations of the form $P \Rightarrow_{\{E, C, A\}} Q$, a prima facie reason similar to the defeasible modus ponens rule mentioned above is needed. One version of this reason is the qualitative version of Pollock's prima facie reason for the statistical syllogism as given in Bex et al. (2003). Because of the importance of this reason, it is defined separately from the other domain-specific prima facie reasons.

**Definition 5.1.2 (Defeasible Modus Ponens/DMP)**

*$\varphi$ and $g_i(\vec{c})$: $\varphi \Rightarrow \psi$* is a prima facie reason for $\psi$

*$\neg valid(g_i)$ undercuts DMP on all ground instances of $g_i$.*
*$exc(g_i(\vec{c}))$ undercuts DMP on $g_i(\vec{c})$.*

As for the notation in this and other definitions of prima facie reasons, if reason $r$ says that *$\varphi$ is a prima facie reason for $\psi$*, then *$\chi$ undercuts $r$* is shorthand for *$\chi$ is a prima facie reason for $\neg$($\varphi$ is a prima facie reason for $\psi$)*. This presupposes that metalinguistic prima facie reasons can somehow be expressed in the object language. For this Pollock (1995) introduces the braces $\lceil$ and $\rceil$ which denote the "objectification" of an expression in the meta-language. Now the metalinguistic rule '$\varphi \gg \psi$', which stands for "$\varphi$ is a prima facie reason for $\psi$" can be transformed into an expression in the object language, that is, a wff: $\lceil\{\varphi_1,\ldots,\varphi_n\} \gg \psi\rceil$. In this way, *$\chi$ undercuts $r$* is shorthand for *$\chi$ is a prima facie reason for $\lceil\neg r\rceil$*.

The prima facie reason for defeasible modus ponens allows for inference of a conclusion from a defeasible generalization and its consequent. For example, if we believe that Haaknat decided to rob the supermarket and that he is the kind of person who acts on his decisions, we can infer that he actually robbed the supermarket:

(1) Haaknat decides to rob the supermarket
(2) $g_c$(Haaknat, supermarket): Haaknat decides to rob supermarket $\Rightarrow_C$ Haaknat robs supermarket
(3) Haaknat robs the supermarket (***1, 2, DMP***)

The above inference involves causal reasoning. An example of evidential reasoning is where it is inferred from evidence that the witness actually saw Haaknat:

(1) $e_1$: witness $W_1$ saw someone who looked like Haaknat
(2) $g_L(w, x)$: witness $w$ saw someone who looked like $x \Rightarrow_E w$ saw $x$
(3) witness $W_1$ saw Haaknat (***1, 2, DMP***)

Technically, the generalization on line 2 of this evidential inference should be $g_L(W_1$, Haaknat). However, often the general version of the generalization (that acts as a scheme for the grounded version) is given in examples, so that it is clear exactly which version has been used in the reasoning. In the above reasoning, for example, it would also have been possible to use the more specific versions $g_L(w, H)$, $g_L(W_1, y)$ and $g_L(W_1, H)$, as these all act as schemes for the ground instance $g_L(W_1, H)$. This distinction is important as each of the versions of $g_L$ represents a different belief: if we believe everyone can make a judgement about whether (s)he saw Haaknat, we might accept $g_L(w, \mathsf{H})$ and if we believe only $W_1$ can make an accurate judgement about this we might accept Haaknat $g_L(\mathsf{W}_1, \mathsf{H})$. When reasoning with generalizations, it can be important to explicitly indicate which generalization has been used in the inference, as the various versions of $g_L$ can be attacked in different ways (see below). In the above argument we accept that anyone can make judgements about having seen any other person.

Just as in Pollock's work, the possible undercutters are explicitly given with each prima facie reason. Notice that for the undercutters, a distinction is made between on the one hand arguing that a generalization is not valid in general and on the other hand arguing there is only an exception in the particular case, that is, the distinction between attack number 3 and 4 on p. 51. The technical difference here is that exceptions can only be given for particular ground instances of generalizations while it is possible to attack the validity of generalization schemes (i.e. generalizations that contain only variables and thus act as a scheme for their ground instances). For example, with $\neg valid(g_L)$ it is argued that it is never the case that if anyone sees someone who looks like another person then this supports that they saw this other person. By contrast, an exception can only be given for the ground instance, viz: $exc(g_L(\mathsf{W}_1, \mathsf{H}))$. This exception does not deny that, in general, when you see someone looking like person $x$ you saw $x$ but rather that in this case there is an exception to the rule: for example, this particular witness cannot accurately recognize faces or Haaknat has a very common appearance.

The current distinction between validity of generalizations and exceptions to generalizations was also made in earlier work. In Hage and Verheij's Reason-Based Logic (Hage, 1996),[4] a rule (roughly equivalent to my generalization) can be valid or excluded: Valid(rule) or Excluded(rule). Here, the rule's validity or its exclusion are properties of the rule. In order for a conclusion to be drawn from a rule and its conditions, the rule should be valid and not excluded. Prakken (1997) basically allows for a defeasible rule $r$ (again equivalent to my generalization) to be undercut by arguing for $exc(r)$ or $\neg valid(r)$. In his DefLog system, Verheij (2003a) does not explicitly distinguish between the validity of a generalization and exceptions to a generalization. However, in (Verheij, 2005a) he distinguishes between attack on a general version of a generalization containing variables and a specific ground instance of a generalization. Hence, it is possible to distinguish between an attack on "if witness $w$ saw someone who looked like $x$ then $w$ saw $x$" and an attack on "if witness $W_1$ saw someone who looked like Haaknat then $W_1$ saw Haaknat".

In the current formalisation, it is assumed by default that there is no exception to a generalization and an explicit reason has to be given for an exception if one wants to undercut DMP. This corresponds to Prakken and Hage and Verheij's ideas on exceptions. In the current logic, it is also assumed by default that a generalization is valid. In Reason-Based Logic, a rule's validity has to be explicitly established before it is possible to reason with the rule. Prakken argues that both ways are

[4]Hage's theory of rules and reasons was initiated by Hage and further formally developed in cooperation with Verheij (cf. Verheij, 1996; Verheij et al., 1998).

possible: either assume by default that rules are valid or require that an explicit argument for *valid*($r$) has to be given before rule $r$ can be used. In order to correctly reason about validity and generalizations, we should be able to infer a generalization $g$ from *valid*($g$). In this way, providing an argument for *valid*($g$) makes it possible to use the generalization $g$ in further reasoning. Sartor (2008) has argued that the derivation of $g$ from *valid*($g$) proceeds through a separate inference rule.[5] In this book the following conclusive reason from Definition 5.1.3 on p. 108 will be used.

**Definition 5.1.3 (Validity of Generalizations)**

*valid*($g_i$) is a conclusive reason for *all ground instances of* $g_i(\vec{t})$: $\varphi \Rightarrow \psi$.

Further examples of reasoning about the validity of generalizations will be given after the proper definitions of arguments and attack and defeat between arguments have been given. First, the issue as to what kinds of prima facie reasons should be included in the current logic for evidential reasoning has to be further resolved.

The defeasible object language and the DMP reason allow for all types of reasoning in the hybrid theory. Consequently, the *defeasible logic* can be defined as follows:

**Definition 5.1.4 (Defeasible Logic)**

*Let $\mathcal{L}$ be a first-order language and let $\mathcal{R}$ be any sound and complete set of conclusive reasons defined over $\mathcal{L}$. The current* defeasible logic *is a tuple $\mathcal{D} = (\mathcal{L}_\mathrm{d}, \mathcal{R}_\mathrm{d})$, where $\mathcal{L}_\mathrm{d}$ is the defeasible language as per Definition 5.1.1 and $\mathcal{R}_\mathrm{d}$ is $\mathcal{R}$ extended with the prima facie reason for defeasible modus ponens* (*Definition 5.1.2*) *and the reason for validity of generalizations* (*Definition 5.1.3*).

This logic is a combination of an object-language and metalinguistic inference rules. This logic can be further extended with, for example, new prima facie reasons; however, the basic logic as defined here will be implicitly assumed in the rest of this chapter.

It should be noted that the current defeasible logic has no traditional model-theoretic semantics, where a proposition logically follows (from given premises) if and only if it is true in all models (in which the premises obtain). Following what is common in nonmonotonic systems such as those presented by, for example, Reiter (1980); Pollock (1995); Prakken (1997), in the current approach it is assumed that rules or generalizations are somehow given (or can be proposed) and that the meaning of these generalizations lies not in their correspondence with some model of the world but rather in their role in dialectical inquiry.[6]

[5] Sartor argues that the legality of a (legal) norm ($n$: $p \Rightarrow q$ is legal), modelled in the logic of Prakken and Sartor (Prakken and Sartor, 1996, see p. 4), allows for the derivation of the norm $n$ itself.

[6] For more on the discussion of model-theoretic semantics for nonmonotonic logics, see Prakken and Vreeswijk, 2002.

## 5.2 A Formal Theory for Argumentation

In this section the formal, logical aspects of the argumentative part of the hybrid theory will be presented. Because the argument-based part of the hybrid theory is not directly influenced by the story-based part, standard definitions of a formal argumentation system can be used to define this argumentative part. In Section 3.1, the concepts that play a part in reasoning with arguments were discussed in an informal way and it was shown how these notions of argument and defeat correspond to ideas from authors in the field of legal reasoning with evidence. It will therefore in this chapter be assumed that the reader is familiar with notions such as, for example, argument, defeat and reinstatement.

According to Prakken and Vreeswijk (2002) a formal argumentation system, also called a logic for defeasible argumentation, has five basic elements: an underlying logic, a definition of an argument, definitions of attack and defeat among arguments and a definition of the dialectical status of arguments. This work is a synthesis of various ideas and definitions from the literature on formal argumentation; it combines Prakken's (1997)[7] basic conceptions of logical language, defeasible logic and arguments, Pollock's (1995) ideas on prima facie reasons and types of attack on arguments and Dung's (1995) argumentation-theoretic semantics for determining the dialectical status of arguments. Many of the ideas proposed here have also been discussed in earlier work (e.g. Bex et al., 2003; Bex and Prakken, 2004, 2008; Bex et al., 2007a).

### *5.2.1 A Defeasible Logic for Evidential Arguments*

The underlying logic of the argumentation system is basically the defeasible logic for the hybrid system as defined in Section 5.1. However, some additions can be made in the specific context of evidential reasoning. Recall that it was argued that in the current context it would be sensible to have some domain-specific inference rules that correspond to oft-used generalizations. Defeasible modus ponens allows for flexible reasoning with all kinds of generalizations, but fixed and domain-specific prima facie reasons also play an important part in evidential reasoning. In Definition 5.2.1 on p. 110 list of prima facie reasons based on this discussion is given; this list includes prima facie rules based on general epistemic principles and prima facie reasons for reasoning with the various types of evidence. The reasons and the accompanying discussion is for a large part based on earlier work (Bex et al., 2003) and each of the reasons can be viewed as an argumentation scheme. For each reason, some standard undercutters that attack the inference are given; one or two

[7] Prakken uses the argumentation system from his thesis in much of his work and has proposed extensions of his basic work together with Sartor (e.g. Prakken and Sartor, 1997). In this thesis, I will mostly refer to Prakken's original system.

general undercutters are given because determining all undercutters requires extensive knowledge and discussion of theories of perception, memory, statistical analysis and so on. This is not the aim of this book and therefore the undercutters given here are not meant to be exhaustive but only serve as an indication of the kinds of attacks that are possible.

**Definition 5.2.1 (Evidential Reasons)**

(1) **Perception**
*Person* $a_i$ *having a percept with content* $\varphi$ is a prima facie reason for person $a_j$ to believe $\varphi$ (where $a_i$ and $a_j$ may or may not be identical).

*The circumstances are such that* $a_i$ *having a percept with content* $\varphi$ *is not a reliable indicator of* $\varphi$ undercuts the perception reason.

(2) **Memory**
*Person* $a_i$ *recalling* $\varphi$ is a prima facie reason for person $a_j$ to believe $\varphi$ (where $a_i$ and $a_j$ may or may not be identical).

$\varphi$ *was originally based on beliefs of which one is false* or $a_i$ *incorrectly remembers* $\varphi$ undercuts the memory reason.

(3) **Witness Testimony**
*Witness w says "*$\varphi$*"* is a prima facie reason for believing $\varphi$.

*Witness w's veracity is questionable* undercuts the witness testimony reason.

*Witness w's objectivity is questionable* undercuts the witness testimony reason.

*Witness w's observational sensitivity is questionable* undercuts the witness testimony reason.

(4) **Expert Testimony**
*expert e says "*$\varphi$*"* and $\varphi$ *is within domain d* and *e is expert in domain d* is a prima facie reason to believe $\varphi$.

*Expert e's veracity is questionable* or *Statement* $\varphi$ *is not based on backup evidence* undercuts the expert testimony reason.

(5) **Documentary Evidence**
*Document d contains information* $\varphi$ is a prima facie reason to believe $\varphi$.

*Document d's authenticity is questionable* undercuts the documentary evidence reason.

(6) **General Knowledge**
*It is general knowledge that "*$\varphi$*"* is a prima facie reason to believe $\varphi$.

$\varphi$ *is infected by prejudice or value judgement* undercuts the general knowledge reason.

Reasons 1 and 2 are the slightly adapted versions of Pollock's prima facie reasons for perception and memory (Pollock, 1987, p. 35, 1995, pp. 52–57) as proposed by Bex and colleagues (2003, p. 138). The main difference with Pollock's original reasons is that the original reasons only allow for the inference of beliefs from one's own observations or memories. This allows police officers to infer beliefs from their observations in the field, jurors to infer beliefs from their observations of tangible evidence in court and judges to reason from their own observations.[8] However, in reasoning with evidence, observations or memories of other people may also be used by persons to form beliefs; for example, a witness telling the jurors that he remembers a certain event allows the jurors to infer that this event happened. The above rule allows for the inference of beliefs from both one's own and someone else's perceptions.

As for the undercutters of reason 1 and 2, the perception undercutter and the first undercutter of the memory reason are literally adopted from (Pollock, 1987). The second undercutter for the memory rule has been slightly adapted and basically stands for the situation where someone remembers something that not really happened. For example, the person could have dreamt that a certain event happened. Another example of such misremembering is what Wagenaar and colleagues (1993, pp. 122–123) call the problem of placement: after a witness saw a possible suspect in the newspaper or on television, the witness might think he saw the suspect at the scene of the crime, when in reality the witness only recognizes the suspect from the news.

Recall that each type of evidential data (e.g. testimonial, tangible) essentially has its own associated generalization which is used to draw conclusions from the evidential data. Reasons 3 and 4 concern two types of testimonies. First consider reason 3 for witness testimony. Since testimonial evidence plays such an important part in reasoning with criminal evidence, it can be assumed that testimonies are a valid source of evidence. Anderson and colleagues (2005) also seem to implicitly accept the witness testimony generalization as being true by default; in their example of a Wigmore chart, they leave the witness testimony generalization implicit while explicitly mentioning other, less typical generalizations in their line of reasoning.

Notice that reason 3 can be undercut when there are doubts about the witness' veracity, objectivity and observational sensitivity. Essentially, only an undercutter for veracity would be needed here: a witness will usually talk about his or her past observations, so usually it will be possible to interpret $\varphi$ as "I recall that I observed $\psi$". On this account, arguments using witness testimonies apply a chain of three prima facie reasons: first the witness scheme is used to infer "I recall I observed $\psi$", then the memory scheme provides "I observed $\psi$" and finally the perception scheme yields $\psi$. Thus lack of objectivity can be handled by undercutters of both memory and perception, and defects in observational sensitivity by undercutters of perception. However, to reduce the complexity of the inferences in the logic, lack of objectivity and observational sensitivity also undercuts the witness testimony reason. Note that it is still possible to construct the above-mentioned

[8] art. 339 par. 1 sub. 1 and art. 340 DPC, which determine that a judge's observations are legally valid evidence.

chain and thus change the witness testimony reason into three separate reasons (see Fig. 3.13, p. 52).

Reason 4 for reasoning from expert testimony shows how the argumentation scheme as depicted in Section 3.1.3 can be modelled as a prima facie reason and how some of the critical questions can point to undercutters of this prima facie reason. The two undercutters presented here correspond to critical questions 4 and 6, respectively. The other critical questions do not directly point to undercutters but point to other ways in which an argument based on the expert testimony prima facie rule can be attacked. The first critical question is a matter of making an inference based on an expert testimony stronger, the fifth refers to rebutting applications of the same prima facie reason and the second and third question seem to be challenges of the premises of the above prima facie rule. This idea of formalizing argumentation schemes and their accompanying critical questions has gained much interest in the past few years. Following Verheij's (2003b) methodology for analysing argumentation schemes, Verheij himself and Gordon et al. (2007) incorporated argumentation schemes in a formal framework.

The two testimonial reasons serve as examples and it is possible to model other kinds of testimonies (e.g. police testimony, suspect testimony) as separate reasons with different undercutters. For current purposes it will be assumed that such other testimonies are simply variants of the witness testimony prima facie rule as defined above, as the general structure and idea of these prima facie rules for other testimonies is the same. More specific undercutters for inferences from, for example, suspects' statements could be defined but, as was already mentioned, the undercutters given here are not meant to be exhaustive but only as an indication of the kinds of attacks that are possible.

The prima facie reasons for testimonies and perception allow for reasoning with the two types of evidence, testimonial and tangible, particularly tangible evidence like objects or traces. Furthermore, reason 5 allows for reasoning with documentary evidence, a type of reasoning that dominates Dutch criminal proceedings and reason 6 allows for reasoning from general knowledge (see Section 2.1.3). Reasoning from personal experience is also possible using the memory and perception rules.

Given these additional prima facie rules, the extended logic for evidential reasoning can now be defined as follows:

**Definition 5.2.2 (Defeasible Logic for Evidential Reasoning)**

*Let $\mathcal{L}$ be a first-order language and let $\mathcal{R}$ be any sound and complete set of conclusive reasons defined over $\mathcal{L}$. The current* defeasible logic for evidential reasoning *is a tuple $\mathcal{E} = (\mathcal{L}_\mathrm{d}, \mathcal{R}_\mathrm{E})$, where $\mathcal{L}_\mathrm{d}$ is the defeasible language as per Definition 5.1.1 and $\mathcal{R}_\mathrm{E}$ is $\mathcal{R}$ extended with the reason for defeasible modus ponens* (*Definition 5.1.2*) *and the evidential reasons (Definition 5.2.1).*

In Section 5.1 it was shown how such a logic allows us to infer conclusions from other propositions and thus build arguments. Arguments are based on some input information consisting of conditional generalizations, evidential data and defeasible data. This input forms the basis of reasoning with arguments:

**Definition 5.2.3 (Evidential Theory)**

*An* evidential theory *$ET$ is a set* $G_E \cup I_E$, *where*

- $G_E$ *is a set of* evidential generalizations.
- $I_E$ *is a consistent set of* evidential data.

The set of evidential data $I_E$ contains the facts of the case that cannot be sensibly denied because the reasoner has witnessed them through autoptic proference.[9] The set $G_E$ contains only evidential generalizations of both the conditional and non-conditional form (causal and abstraction generalizations are not part of the argumentative part of the theory). This set is essentially the current stock of knowledge and it consists of generalizations, general knowledge, knowledge from personal experience of the reasoner, opinions or ideas which may be called into question or actively denied. Unless mentioned otherwise, the discussion below assumes an arbitrary but fixed theory *ET*.

### *5.2.2 Evidential Arguments*

Given a particular evidential theory and the evidential logic $\mathcal{E}$, arguments can be constructed as follows:

**Definition 5.2.4 (Evidential Arguments)**

*An argument based on an evidential theory ET is a finite sequence* $[\varphi_1,...,\varphi_n]$, *where* $n > 0$, *such that for all* $\varphi_i$ $(1 \leq i \leq n)$:

- $\varphi_i \in ET$; *or*
- *There exists a (conclusive or prima facie) reason in* $\mathcal{R}_\mathsf{E}$ *such that* $\varphi_1,...,\varphi_m \in \{\varphi_1,...,\varphi_{i\text{-}1}\}$ *is a reason for* $\varphi_i$.

*For any evidential theory ET the set of all arguments based on ET is denoted by* $Args_{ET}$.

$\varphi$ *is* $\mathcal{E}$*-derivable from ET (denoted as* $ET \vdash_\mathsf{E} \varphi$*) iff there exists an argument based on ET which has* $\varphi$ *as its last element.*

The elements of the sequence are also called *lines of argument*. According to the above definition, a line of argument is a proposition from the input information *ET* or is derived from preceding lines of argument by the application of some reason. The above definition is adapted from Prakken (1997, pp. 154–155), calls a derivation of the above form a *default deduction*. The similarity between arguments and proofs is further illustrated by the defeasible logical consequence relation $\vdash_\mathsf{E}$ (Prakken renders this as |~).

[9] Formally, consistency of the set $I_E$ can only be defined if the pieces of evidence are not labelled with their names. However, here it will be simply assumed that consistency is determined on the basis of the propositional content of the evidential data, ignoring the names.

Notice that arguments can also be defined with more inherent features; for example, we could require that arguments are consistent or, as in Prakken's definition, all defeasible rules (generalizations) are used in the argument itself and that no element occurs more than once. For now, an argument will be defined as above. This definition is similar to Pollock's definition (1995, pp. 89–90): Pollock regards the two conditions as argument formation rules, where the first one is called "Input" (inserting a proposition from the input set into an argument) and the second one is called "Reason" (applying a reason to a line of argument). Recall from Section 3.1.2 that an argument can have an associated strength. In the rest of this book, however, the strength of arguments plays no important role so it will be left implicit.

As an example of an argument that uses multiple reasons, consider the following:

*Argument* (*H in car*)

(1) $e_{\mathrm{w1}}$: $\mathrm{W}_1$ testifies that "I saw someone looking like Haaknat get into the car". (***$I_E$***)
(2) witness $\mathrm{W}_1$ saw someone who looked like Haaknat get into the car. (***1, testimony***)
(3) $g_{\mathrm{L}}$: a witness *w* saw someone who looked like $x \Rightarrow_{\mathrm{E}} w$ saw *x*. (***$G_E$***)
(4) $\mathrm{W}_1$ saw Haaknat get into the car. (***2, 3, DMP***)
(5) Haaknat got into the car. (***4, perception***)

Note that here the lines of argument have been numbered and that at the end of each line of argument it is noted where the proposition stems from (i.e. input or inferred from some other lines in the argument). Furthermore, for simplicity's sake the name of the generalization has been shortened by leaving out the tuple denoting the terms. Arguments can also be given for the validity of generalizations.

*Argument* (*validity* $g_{\mathrm{L}}$)

(1) $g_{\mathrm{gkL}}$: $\Rightarrow_{\mathrm{E}}$ It is general knowledge that "witnesses can accurately recognize people". (***$G_E$***)
(2) witnesses can reasonably accurately recognize people. (***1,Gen. Knowl.***)
(3) $g_{\mathrm{Lvalid}}$: witnesses can reasonably accurately recognize people $\Rightarrow_{\mathrm{E}}$ *valid*($g_{\mathrm{L}}$) (***$G_E$***)
(4) *valid*($g_{\mathrm{L}}$) (***2, 3, DMP***)
(5) $g_{\mathrm{L}}$: a witness W1 saw someone who looked like Haaknat $\Rightarrow_{\mathrm{E}}$ $\mathrm{W}_1$ saw Haaknat. (***4, conclusive***)

Here line 1 is a non-conditional generalization from $G_E$, a piece of general knowledge. The inference from line 4 to 5 is warranted by the conclusive rule from Definition 5.1.3.

As was argued in Chapter 1, the formal theory as described in this section serves as the basis for a view on sense-making and visualisation of scenarios. In the software tool AVERS, van den Braak (2010) uses the visual AIF argument language (Rahwan and Reed, 2009), which renders arguments in a similar way to the arguments in Chapter 3. Bex et al. (2010) have shown that a logic not too dissimilar to the current one can be formally connected to these diagrams. In this thesis, this connection between logic and diagram will be shown through examples. For example, the above two arguments can also be rendered as graphs (Fig. 5.1).

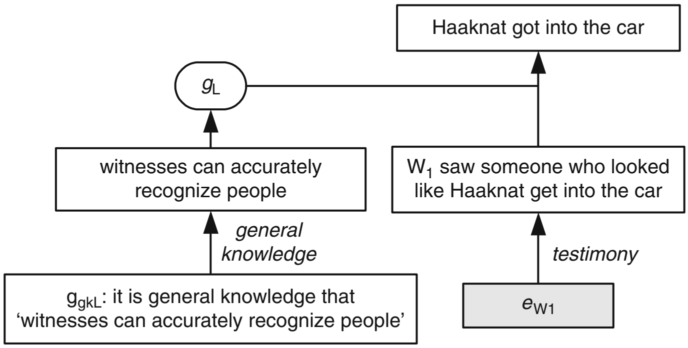



**Fig. 5.1** Graphical representation of arguments

Here, the arrows stand for evidential inferences and the boxes stand for lines of argument; a gray box is a piece of evidence and a white box is a piece of general knowledge from $G_E$ (such as $g_{gkL}$). Rounded boxes denote conditional generalizations that warrant an inference and prima facie reasons are other than *DMP* are mentioned next to the arrow. Argument (*validity* $g_L$) has been simplified: the inference from line 2 and 3 to line 4 (with the generalization $g_{Lvalid}$) and the inference from line 4 to line 5 has been summarized as a single inference. When there is no risk of confusion, other arguments will also for clarity be summarized in this way.

The following auxiliary notions are related to the notion of arguments and will be used in this chapter:

**Definition 5.2.5 (Premises, conclusions, input and subarguments)**

*For any argument A and A'*

(1) *The set of* premises *of A is* $Prem(A) = \{\varphi \in A \mid \varphi \in ET\}$;
(2) *The set of* conclusions *of A is* $Conc(A) = \{\varphi \in A \mid \varphi \notin ET\}$;
(3) *The set of* generalizations *in A is* $Gens(A) = \{g_i \in A \mid g_i \in G_E\}$;
(4) *The set of* evidence *in A is* $Evidence(A) = \{e_i \in A \mid e_i \in I_E\}$;
(5) *An argument A' is a (proper)* subargument *of A iff A' is a (proper) subsequence of A.*
(6) *An argument A is* strict *if it does not contain a generalization from* $G_E$; $\varphi_i$ *is* defeasible *otherwise.*

Number 1 and 2 say that propositions that follow from the input information are considered premises and the other propositions are conclusions. In argument (*H in car*), the propositions in line 1 and 4 are premises and the rest are conclusions. Numbers 3 and 4 define the sets of (non-defeasible) evidence and defeasible generalizations (the information from stock of knowledge) that is used in a particular argument. Number 5 defines subarguments and number 6 defines the dichotomy between *strict arguments*, consisting only of evidential data (which cannot be denied) and conclusive rules (which do not make arguments defeasible) and *defeasible arguments*, which incorporate assumptions from the stock of knowledge or use prima facie reasons. The notions of premises, conclusions, subarguments and strict and defeasible arguments were taken from Prakken (1997). No minimality requirement on arguments has been defined as for current purposes this is not necessary. Note that for simplicity, often conclusions that are inferred through conclusive reasons are simply said to follow from the argument. For example, from argument (*H in car*) it can be inferred through conclusive reasons that ¬¬(Haaknat got into the car) or $W_1$ saw Haaknat get into the car ∧ Haaknat got into the car. Technically, these are both separate arguments with (*H in car*) as their subarguments but in examples I will simply say that ¬¬(Haaknat got into the car) and $W_1$ saw Haaknat get into the car ∧ Haaknat got into the car are conclusions of (*H in car*).

### *5.2.3 Attacking Arguments*

Now the notion of attack between arguments can be formally defined. As for notation, the definition includes the complement of a formula $\varphi$, denoted as $\overline{\varphi}$. This complement of $\varphi$ is $\neg\psi$ if $\varphi = \psi$ and $\psi$ if $\varphi = \neg\psi$.

**Definition 5.2.6 (Attack Among Arguments)**

*Given two arguments A, B ∈ $Args_{ET}$:*

(1) $\varphi \in A$ rebuts $\psi \in B$ *iff* $\varphi = \overline{\psi}$ *and* $\psi$ *is the conclusion of a defeasible argument.*
(2) $\ulcorner\neg(\{\varphi_1,\ldots,\varphi_n\} \gg \psi)\urcorner \in A$ *undercuts* $\psi \in B$ *iff* $\psi$ *is obtained from some earlier lines* $\varphi_1,\ldots,\varphi_n$ *in B by the application of a prima facie reason.*

*An argument A attacks an argument B iff a line of argument in A either rebuts or undercuts a line of argument in B.*

An argument *A* rebuts another argument *B* if *A* negates some defeasible input information or a generalization used in *B* or if it negates some defeasible conclusion that has been inferred through a prima facie reason. Note that if both $\varphi$ and $\psi$ are defeasible, then the two arguments attack each other and rebutting attacks can thus be symmetrical. An argument *A* undercuts another argument *B* if it negates some defeasible inference in *B*. Definition 5.2.6 essentially provides a combination of assumption-based argumentation (e.g. Bondarenko et al., 1997; Poole, 1988; Prakken and Sartor, 1997), where the nonmonotonicity lies in the fact that the arguments are based on defeasible assumptions which can be contradicted, and inference-based argumentation (e.g. Pollock, 1995), where the nonmonotonicity is modelled by allowing attacks on the defeasible inferences. A similar combination can also be found in Verheij's DefLog (Verheij, 2003a).

Note that attacking an argument does not mean that the argument that is attacked is also defeated. First, some examples of attacks between arguments are given. Argument (*H in car*) can be rebut by the following argument:

*Argument* (*H not in car*)

(1) $e_{\mathrm{H}}$: Haaknat testifies that "I never got into that car". (***$I_E$***)
(2) Haaknat never got into the car. (***1, testimony***)
(3) ¬(Haaknat got into the car). (***2, conclusive***)

Notice that here the exact conclusive reason that is used to infer line 3 from line 2 is not mentioned: it is clear that if Haaknat never got into the car then he did not get into the car when the witness thought he saw him. When there is no danger of confusion, such simple inferences will be considered conclusive and left implicit.

The argument (*H in car*) can also be undercut by attacking an inference. For example, the *testimony* prima facie reason can be undercut by providing an argument against $W_1$'s veracity, or the perception reason can be undercut by arguing that $W_1$ could not have seen who got into the car because it was a very foggy day. However, for now it might be interesting to provide examples that connect to the discussion on the sources, exceptions and validity of generalizations (see Section 3.1.3). Argument (*H in car*) uses one generalization, namely $g_L(w, x)$. Now, an argument for an exception $exc(g_L(W_1, H))$ undercuts the application of DMP to $g_L$ in (*H in car*):

*Argument* (*common appearance*)

(1) $g_{ca1}$: $\Rightarrow_E$ Haaknat has a common appearance. (***G<sub>E</sub>***)
(2) $g_{ca2}$: $x$ has a common appearance $\Rightarrow_E$ $exc(g_L(W1, H))$. ($\boldsymbol{G_E}$)
(3) $exc(g_L(W_1, H))$. (***1, 2, DMP***)
(4) $\neg$ ⌈witness W1 saw someone who looked like H get into the car, $g_L$(W1, H):witness W1 saw someone who looked like H $\Rightarrow_E$ W1 saw H $\gg$ $W_1$ saw H get into the car⌉ (***3, Definition 5.1.2***)

Note that line 4 will normally not be explicitly included in an argument. It is also possible to attack the generalization's validity:

*Argument* (*invalidity* $g_L$)

(1) $e_{e1}$: expert *E*2 says "people are notoriously bad at recognizing others". ($\boldsymbol{I_E}$)
(2) people are bad at recognizing others. (***1, expert***)
(3) $g_{br}$: people are bad at recognizing others $\Rightarrow_E$ $\neg valid(g_L)$. ($G_E$)
(4) $\neg valid(g_L)$. (***3, DMP***)

The above argument (in which for simplicity not all the premises of the expert testimony reason are given) argues that it is usually not the case that "if a witness sees someone who looks like person $x$ then the witness saw $x$" because people are bad at accurately making judgements about whether the two persons they saw were the same. In argument (*H in car*), the validity of $g_L$ was assumed by default. By giving argument (*invalidity* $g_L$), this validity is now called into question.

Note that in the above arguments, it is possible to undercut the inference of the exception $exc(g_L(W1, H))$ or the invalidity $\neg valid(g_L)$. For example, it can be argued that "the fact that a person has a common appearance does not influence other people's ability to recognize him'; this is undercuts the DMP inference from line 2 to line 3 in argument (*common appearance*) because it can essentially be seen as a reason for $\neg valid(g_{ca})$. In this way, it is possible to talk about the invalidity of exceptions and similarly about exceptions to exceptions.

Attacks can also be rendered graphically by using an arrow with a round arrowhead (see Fig. 5.2). Here the conclusion of (*H not in car*) rebuts the conclusion of (*H in car*); the fact that rebuttal is symmetrical is represented by the two round

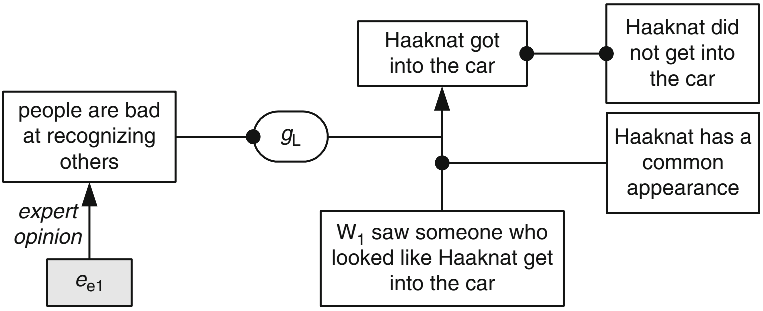



**Fig. 5.2** Graphical representation attacks between arguments

arrowheads on the same line. Whilst both (*common appearance*) and (*invalidity* $g_L$) undercut the inference with the generalization $g_L$, the argument for $exc(g_L)$, (*common appearance*), is rendered as attacking the inference arrow itself and the argument for $\neg valid(g_L)$, (*invalidity* $g_L$), is rendered as attacking the generalization, so that these two attacks are properly distinguished. Notice that, as in Fig. 5.1, some inferences have not been explicitly rendered in Fig. 5.2. For example, argument (*invalidity* $g_L$) has been rendered as one inference.

### 5.2.4 Defeat and the Status of Arguments

Given the above definitions of an evidential theory, arguments and defeat, an evidential argumentation theory can be defined:

**Definition 5.2.7 (Evidential Argumentation Theory)**

*An evidential argumentation theory is a tuple $AT = (ET, Args, Attack)$ where:*

- *$ET$ is an evidential theory;*
- *$Args$ ($\subseteq Args_{ET}$) is a set of arguments constructed on the basis of $ET$;*
- *$Attack$ ($\subseteq Args \times Args$) is a binary relation containing the pairs of arguments that attack each other.*

This evidential argumentation theory is a specific instance of an argument framework in the style of (Dung, 1995), which is specifically based on an evidential theory. Such a theory is a collection of arguments and their attack relations based on a specific *ET*. In other words, the evidential theory can be seen as the input of a specific argumentation theory. The rest of this chapter always assumes some arbitrary but fixed evidential argumentation theory, unless explicitly mentioned.

Now a definition of defeat can be given:

**Definition 5.2.8 (Defeat Between Arguments)**

*An argument $A$* defeats *an argument $B$ iff $(A, B) \in Attack$.*

Normally, we say that an argument *A* defeats another argument *B* if and only if *A successfully attacks B*. Because currently no notion of strength or preference of arguments is defined, the notions of attack and defeat have simply been equated. In Section 3.1.2 several ways of resolving conflicts between arguments in a detailed way were discussed; for current purposes, however, the above definition is sufficient.

Given a definition of defeat, the dialectical status of the arguments can be determined. First, we should determine which arguments can defend themselves against incoming attacks. An evidential argumentation theory *AT* is based on an argumentation framework as defined by Dung (1995), so we can use Dung's definitions.

**Definition 5.2.9 (Acceptable Arguments and Admissible Set)**

*Given some evidential argumentation theory AT = (ET, Args, Attack, Pref)*

- *A set of arguments SA is said to be* conflict-free *if no argument in SA is defeated by another argument in SA.*
- *An argument A is* acceptable *w.r.t. a set of arguments SA iff for each argument B ∈ Args it is the case that if B defeats A then B is defeated by a member of SA.*
- *A conflict-free set of arguments SA is* admissible *iff each argument in SA is acceptable w.r.t. S.*

This definition captures the notions of indirect support and reinstatement in an intuitive way. Consider a theory $AT_1$ with a set of arguments {(*H in car*), (*H not in car*)} and their attack relations {((*H in car*), (*H not in car*)), ((*H not in car*), (*H in car*))}. The admissible sets are Ø, {(*H in car*)}, {(*H not in car*)}. Now consider theory $AT_2$, where argument (*common appearance*) and its corresponding attack relation ((*common appearance*), (*H in car*)) are added to their respective sets in $AT_1$; for $AT_2$, the admissible sets are Ø, {(*H not in car*)} and {(*H not in car*), (*common appearance*)}.

The notion of admissible sets allows for several definitions of extensions of an (evidential) argumentation theory. Such an extension is a set of arguments that defends itself; the exact composition of the set is determined by the type of extension. For example, Dung defines various kinds of extensions (i.e. preferred, complete, stable, grounded). Some of these extensions present a more sceptical point of view than others; for example, the grounded extension of the theory in $AT_1$ is Ø, while the preferred extensions are {(*H in car*)} and {(*H not in car*)}. Basically this means that if we use grounded semantics, nothing can be said about which argument is winning, while with preferred semantics we can say that a choice has to be made between {(*H in car*)} and {(*H not in car*)}. For current purposes, the exact choice of extension is not important.

Given a particular extension, the dialectical status of the arguments can be determined. Recall from Section 3.1.2 that there are three different statuses of arguments: justified, defensible and overruled. Below preferred semantics is used to define these status assignments (the definition is adapted from Prakken and Vreeswijk, 2002).

**Definition 5.2.10 (Status Assignments for Preferred Semantics)**
*Given some AT* = (*ET, Args, Defeat*), *a preferred extension of AT is a maximal* (*w.r.t. set inclusion*) *admissible set of AT.*

– *an argument A is* justified *iff it is an element of all preferred extensions of AT.*
– *an argument A is* defensible *iff it is an element of some preferred extensions of AT.*
– *an argument A is* overruled *iff it is an element of no preferred extension of AT.*

Since for the theory $AT_1$, the preferred extensions are {(*H in car*)} and {(*H not in car*)}, both arguments in the theory are defensible. In the theory $AT_2$, the argument {(*H not in car*)} is reinstated by (*common appearance*) and hence there is only one preferred extension, namely {(*H not in car*), (*common appearance*)}. Consequently, both (*H not in car*) and (*common appearance*) are justified while (*H in car*) is overruled in $AT_2$.

## 5.3 A Formal Theory for Explanatory Stories

In Section 3.2, stories that explain some evidence were presented in an informal way. This section focuses on the formal aspects of explanatory stories. A general theory on which such stories can be based is presented and the way in which they explain explananda is formally defined. The various uses of abstraction links are discussed and ways in which stories can be connected to story schemes are also defined.

In the hybrid theory, the argumentative part directly influences the causal story-based part of the theory. For example, arguments determine which elements can be in the explananda (see p. 138). Furthermore, they are used to reason about the quality of stories: they are used to support and contradict a story with evidential data and to reason about the coherence of a story. The interaction between arguments and stories will be discussed in Section 5.4 in which the hybrid theory will be presented. In this section the main features of the causal story-based part of the hybrid theory will be discussed.

### *5.3.1 A Causal Theory for Explanations*

In the story part of the hybrid theory, hypothesized stories should causally explain the explananda. As in traditional models of abductive model-based reasoning, this *explains* relation between a story and the explananda can be defined through a notion of logical consequence: the explananda should follow from a combination of causal generalizations expressing relations between events and (conjunctions of) literals denoting the initial events which are in the generalizations' antecedents. The logic for defeasible inferences as given in Definition 5.1.4 can essentially also be used for this kind of causal inference. However, the logic should naturally not include the specific reasons for evidential reasoning (e.g. witness testimony). Thus, the causal logic can be seen as a more abstract logic than the evidential logic. It is possible to provide a contextual logic for causal reasoning with, for example, specific prima facie rules for the various types of causal inference. However, an elaboration of such ideas would first require a more thorough exposition and analysis of causality, which will not be presented in this book.

Now, a *causal theory* that is to be used as the basis for stories should contain the explananda and the literals and generalizations from which the explananda may follow:

**Definition 5.3.1 (Causal Theory)**

*A causal theory is a set* $CT = G_{CA} \cup F \cup H$ *where*

- $G_{CA} = G_C \cup G_A$*, where* $G_C$ *is a set of* causal generalizations *and* $G_A$ *is a set of* abstraction generalizations*; and*
- $F$*, the* explananda*, is a consistent set of ground first-order literals; and*
- $H$*, the* hypotheticals*, is the set of all ground literals occurring in the antecedent of some causal generalization in* $G_{CA}$ *and instantiated with some term in* $G_{CA} \cup F$*.*

In order to keep in line with the terminology used in this book, $G_{CA}$ consists of generalizations (instead of "rules" as they are usually called). $G_{CA}$ is a combination of causal generalizations, which allow us to connect the various events in a story, and abstraction generalizations, which allow events in a story to be connected to some abstract version of an event. The hypotheticals $H$ contain the assumed initial events which are in the antecedent of some generalization. They are called hypotheticals so as not to give rise to confusion between a general hypothesis in a case, which can vary in its form and complexity, and a hypothetical, which is always a literal. Finally, in standard definitions of causal-abductive reasoning usually a set of *observations* that has to be explained by the combination of hypotheticals and generalizations is given. In this book these observations are called the *explananda* and the term "observation" is used for conclusions of evidential arguments that can be used to differentiate between explanations.

The causal theory is in a sense analogous to the evidential theory $ET$. The evidential theory supplies the input knowledge (i.e. evidential data and (non-conditional) generalizations) on the basis of which arguments can be constructed. The causal theory $CT$ supplies the input knowledge on the basis of which stories can be constructed, namely assumed events (hypotheticals) and causal generalizations. An important difference is that the causal theory does not contain elements that can be accepted without question, as is the case with the evidence in $I_E \subseteq ET$. In the evidential theory, the generalizations in $G_E$ represent the stock of knowledge; in the causal theory, the combination of the hypotheticals $H$ and the generalizations $G_{CA}$ represent the stock of knowledge, assumed information on the basis of which stories can be constructed.

### *5.3.2 Causal Stories*

Standard accounts of abductive model-based reasoning simulate the abductive inference with classical-logical derivation: some hypothesis $S$ explains an explanandum

$\varphi$ if $\varphi$ is a logical consequence of $S$. In order to define how stories explain explananda in the current logic, first the structure of a story and the way in which propositions can be derived from a story should be defined. A story is a combination of hypotheticals and assumed generalizations, which together allow for the construction of inferences with the Defeasible Modus Ponens reason.

**Definition 5.3.2 (Causal Story)**

*A story S based on a causal theory CT is a finite sequence* $[\varphi_1,\ldots,\varphi_n]$, *where* $n > 0$, *such that for all* $\varphi_i$ $(1 \leq i \leq n)$:

- $\varphi_i \in CT$; *or*
- *According to the Defeasible Modus Ponens reason* $\varphi_1,\ldots,\varphi_m \in \{\varphi_1,\ldots,\varphi_{i-1}\}$ *is a reason for* $\varphi_i$.

*For any causal theory CT the set of all stories based on CT is denoted by* $Stories_{CT}$.

$\varphi$ *is* $\mathcal{C}$*-derivable from CT (denoted as* $CT \vdash_{\mathsf{C}} \varphi$*) iff there exists a story based on CT which contains only causal generalizations and which has* $\varphi$ *as its last element.*

$\varphi$ *is* $\mathcal{A}$*-derivable from CT (denoted as* $CT \vdash_{\mathsf{A}} \varphi$*) iff there exists a story based on CT which contains only abstraction generalizations and which has* $\varphi$ *as its last element.*

$\varphi$ *is* $\mathcal{CA}$*-derivable from CT (denoted as* $CT \vdash_{\mathsf{CA}} \varphi$*) iff there exists a story based on CT which has* $\varphi$ *as its last element.*

Here, a story is a sequence of events that forms a meaningful whole through its causal connectivity. Take the example of a causal inference that was given in Section 5.1, reproduced as a sequence: [Haaknat decides to rob the supermarket, $g_c$: Haaknat decides to rob the supermarket $\Rightarrow_{\mathsf{C}}$ Haaknat robs the supermarket, Haaknat robs the supermarket]. Here, Haaknat decides to rob the supermarket is a hypothetical, $g_c$ is a causal generalization and Haaknat robs the supermarket is derived from this hypothetical and generalization. Note the similarity of a causal story to an evidential argument: both are derivations in their respective defeasible logic. This structural similarity does, however, not mean that stories and arguments are used in the same way. For example, stories cannot attack and defeat each other. Arguments are used to *argue* for a particular conclusion and stories are used to *explain* some explananda. Furthermore, in evidential reasoning arguments are often based on evidential data, whilst stories cannot be directly based on evidence but rather provide hypotheses. In Chapters 3 and 4, the differences between arguments and stories are discussed in detail.

Just as for arguments, auxiliary notions for causal stories can be defined:

**Definition 5.3.3 (Generalizations, Events and Substories)**

*For any story S*

- *The set of* generalizations *in S is* $Gens(S) = \{g_i \in S \mid g_i \in G_{CA}\}$;
- *The set of* events *in S is* $Events(S) = \{$a is a ground literal $\mid$ a occurs in the antecedent or the consequent of some $g_i \in Gens(S)\}$;

– *The set of* initial events *in S is IniEv*(*S*) = {$a \in$ *Events*(*S*) | $a$ does not occur in the consequent of some $g_i \in$ *Gens*(*S*)}*;*
– *A story S' is a (proper)* substory *of S iff S' is a proper subsequence of S.*

Note that as in arguments, technically only grounded instances of generalizations are part of an explanation. However, if the general versions of the generalizations (that act as a scheme for the grounded version) are used, this will be shown in the derivation. The events[10] are all the individual literals in a story. For readability, I will often leave individual events in a story implicit; for example, a story $S = [e_1, e_1 \Rightarrow_C e_2, e_2, e_2 \Rightarrow_C e_3, e_3]$ will be simplified to $[e_1, e_1 \Rightarrow_C e_2, e_2 \Rightarrow_C e_3]$. Here, the events are the instantiated antecedents and consequences of the generalizations. Furthermore, no formal "append" function that adds one story to another will be defined but the combination of two stories *S* and *S'* will simply be written as (*S*, *S*) and a story *S* combined with a wff $\varphi$ is also written as $(S, \varphi)$. So (*S*, $[e_3, e_3 \Rightarrow_C e_4]$) stands for the story $[e_1, e_1 \Rightarrow_C e_2, e_2 \Rightarrow_C e_3, e_3 \Rightarrow_C e_4]$. Similar to arguments, stories can also be represented graphically as was already done in Chapters 3 and 4 (Fig. 5.3). In this figure, the arrows with the open heads represent causal relations. Notice that, for reasons of clarity, case-specific causal generalizations expressing these relations are not always explicitly rendered in a rounded box: in the case of $e_1 \Rightarrow_C e_2$, the generalization is simply expressed by the arrow from $e_1$ to $e_2$.

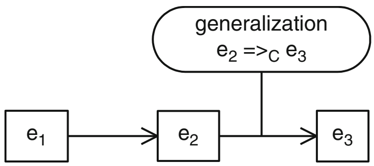



**Fig. 5.3** Graphical representation of stories

### *5.3.3 Stories as Explanations*

In traditional approaches to causal-abductive reasoning with stories, a hypothesis in the form of a network of causal relations explains a set of observations. Formal models of symbolic causal-abductive reasoning (e.g. Console and Torasso, 1991; Bylander et al., 1991; Konolige, 1994) were primarily developed to model medical or system diagnosis.[11] The main idea in these logical models for abductive reasoning is that a set of causal rules *T*, which consists of rules of the form "*cause* → *effect*", together with some hypothesized literals *H* together imply the observations *O*. The causal rules are modelled using material implication, which satisfies the modus ponens rule, and hence *O* can be derived from *H* and *T*. For example, say that $H = \{c\}$, $T = \{c \rightarrow e\}$ and $O = \{e\}$, then $\{c\} \cup \{c \rightarrow e\} \vdash e$.

[10] Strictly speaking, these *events* can be either events or states of affairs.

[11] However, model-based diagnosis should not be equated with abductive reasoning, as there are other types of model-based reasoning (e.g. Bayesian Belief Networks) and other types of symbolic diagnosis (e.g. consistency-based diagnosis). See Lucas, 1997 for an overview.

Because a cause does not always lead to its effect (for example, not all people who rob a supermarket hide), some models (e.g. Console and Torasso, 1991) weaken the causal link by introducing atoms of the form α in the conditions of causal rules so that a causal rule $c \wedge \alpha \rightarrow e$ stands for "*c* may cause *e*". The deduction of the effect *e* can then be blocked by not including α in the hypothesis *H*. Poole (1994) has a different approach to modelling these non-deductive relations. He allows only strict causal rules to be part of *T* and causal rules denoting weaker causal relations can be assumed in *H* but are not part of *T*. So then *c* as well as the rule $c \rightarrow e$ have to be in *H* because if the causal rule is not assumed, it is impossible to derive e.

This book introduces a causal abductive way of reasoning that is slightly different from the traditional models as briefly introduced above. Aside from differences in terminology (i.e. explananda instead of observations or generalizations instead of rules), causal generalizations and abstraction generalizations are modelled using the defeasible implication ⇒. Because the generalizations are now modelled with the defeasible implication no further notation has to be introduced to allow them to represent generalizations of the form "*c* may cause *e*". Furthermore, explaining is defined as there being an explicit story that has the explananda as one of its events. In most logical models of abductive reasoning, some hypotheticals together with some generalizations explain the explananda when these explananda are simply derivable from the hypotheticals and the generalizations (i.e. *hypotheticals* ∪ *generalizations* ⊢ *explananda*). Defining an explanation as a story (i.e. an explicit derivation) as is done in this book ensures that all the intermediate events are also part of the explanation (Definition 5.3.4).

**Definition 5.3.4 (Explanation)**

*Given a CT, a story S is an* explanation *for a set of literals E iff*

(1) $\forall e$: *if* $e \in E$ *then* $e \in Events(S)$*; and*
(2) $S \nvdash_{CA} \bot$*; and*
(3) *S contains no two generalizations with the same consequent.*

The second condition in the above definition, which is standard in models of abductive causal reasoning, ensures that the explanation does not lead to inconsistencies. Notice that this condition effectively models one of the criteria for the quality of a story, namely the consistency criterion (see Section 4.3.1). For the consistency criterion it is not the *extent* to which a story is consistent but rather the question *if* a story is consistent. Explanations that internally contradict themselves can never be fully supported because this would require us to accept contradictory evidence. Therefore, consistency is modelled as a condition on explanations so that stories which are inconsistent are not considered as explanations. The third condition ensures that two different explanations for *F* are really seen as two separate explanations: two stories with the same consequences are considered as alternative stories. In standard approaches to abductive model-based reasoning this is usually enforced by requiring that the explanation is (subset) minimal. However, as will become clear, in the current theory the minimal explanation is not always the best one. Because an explanation as defined in Definition 5.3.4 is a causal story that explains some events, in the rest of this chapter the term "explanation" and the term "story" will be used interchangeably.

As an example, take the basic Haaknat story from Fig. 3.16 (p. 61). For the example, assume that there is only one event which must be explained, namely police find Haaknat in park. The following story explains this explanandum:

$S_p =$ [Haaknat is a drug addict ∧ Haaknat needs money,
$g_{c1}$: Haaknat is a drug addict ∧ Haaknat needs money $\Rightarrow_C$ Haaknat decides to rob supermarket,
$g_{c2}$: Haaknat decides to rob supermarket $\Rightarrow_C$ Haaknat robs supermarket,
$g_{c3}$: Haaknat robs supermarket ∧ police arrive $\Rightarrow_C$ Haaknat gets into car and takes off,
$g_{c4}$: Haaknat gets into car and takes off ∧ Haaknat thinks police are following $\Rightarrow_C$ Haaknat hides in park,
police search in park,
$g_{c5}$: Haaknat hides in park ∧ police search in park $\Rightarrow_C$ police find Haaknat in park]

Note that, because the set of hypotheticals $H$ consists of all literals that are in the antecedent of a causal generalization, there are also other possible stories such as, for example, [Haaknat hides in park, police search in park, $g_{c5}$] or [Haaknat gets into car and takes off, Haaknat thinks police are following, $g_{c4}$, $g_{c5}$]. For now $S_p$ will be used as the main example.

Most of the causal generalizations in the above explanation stand for not much more than some kind of temporal precedence[12] between events and the actual causal relation may be debated. For example, we could argue that money is not a good motivating cause for Haaknat to decide to rob the supermarket. Furthermore, not all possible causal relations have been expressed by a generalization. For example, the police are looking in the park because they believe that the robber is hiding there but no generalization expressing the causal connection between the robbery and the police's search has been given. In the current model, this naïve and ad hoc view on causal reasoning is purposeful, as it allows for the quick and easy construction of stories and thus retains the holistic flavour of reasoning with stories. This interpretation of causality is not a problem, as any objection to doubtful causal links can explicitly be expressed. Furthermore, the criteria on stories in the hybrid theory (see Section 5.4) together with the active dialectical assessment of stories (see Section 5.5) ensure that the stories are causally connected and plausible.

Notice that all the generalizations in the above story are case-specific in that they do not contain variables. Reasoning with stories in this way is relatively simple as it allows one to express a causal relation between two events in the story directly as the related case-specific generalization and no interpretative step is required to change the causal relation into a more general version of the generalization. For

[12] In the current framework, time is not explicitly represented. However, it is assumed that events can only be caused by other events which precede them and thus a sequence of events implicitly assumes temporal relations between the events.

example, if we want to express the causal relation expressed as "Haaknat needs money so he decides to rob supermarket" ($g_{c1}$) in a more general form, we would have to determine whether we believe that "persons who need money will decide to rob some place" ($x$ needs money $\Rightarrow_C$ $x$ decides to rob $y$) or that "persons who need money will decide to rob a supermarket" ($x$ needs money $\Rightarrow_C$ $x$ decides to rob supermarket) or that "if Haaknat needs money he will decide to rob something" (Haaknat needs money $\Rightarrow_C$ Haaknat decides to rob $y$). Whilst in the atomistic, evidential reasoning with arguments these decisions should often be explicitly made, when reasoning with stories we want to be able to quickly build a simple hypothesis. The focus of stories is on the events and when reasoning with stories, the general world knowledge that is used has often not the form of individual causal generalizations but rather the form of a holistic story scheme. This does not mean that the generalizations that express the causal relations between the various events cannot be more general: some or all of the above generalizations can also be given in a more general form.

Above it was shown how arguments, which are used to argue for a particular conclusion, can attack each other. The possibility of attack is what makes reasoning with arguments dialectical. Causal-abductive reasoning takes its dialectical justification from the fact that multiple hypotheses for the explananda can and should be considered in the process of inference to the best explanation. Consider the alternative explanation for why Haaknat was found in the moat on the next page (see also Fig. 4.4, p. 95).

$S_h =$ [argument between Haaknat and Bennie,

$g_{c6}$: argument between Haaknat and Bennie $\Rightarrow_C$ Haaknat feels threatened by Bennie

$g_{c7}$: Haaknat feels threatened by Bennie $\Rightarrow_C$ Haaknat runs away

$g_{c8}$: Haaknat runs away $\Rightarrow_C$ Haaknat runs into police

$g_{c9}$: Haaknat feels threatened by Bennie $\wedge$ Haaknat runs into police $\Rightarrow_C$ Haaknat hides in park

$g_{c5}$: Haaknat hides in park $\wedge$ police search in park $\Rightarrow_C$ police find Haaknat in park]

The explanation $S_h$ now also explains the explanandum police find Haaknat. Note that any story which contains both $g_{c9}$ and $g_{c4}$ is not a proper explanation according to Definition 5.3.4. Here, the two different causes for Haaknat hiding in the park are mutually exclusive in that they cannot both be accepted at the same time. There are now at least two alternative hypotheses explaining the explanandum, namely $S_p$, which argues that Haaknat committed the robbery and hid from the police and $S_h$, which argues that Haaknat hid in the park because he felt threatened by Bennie.

When reasoning with hypothetical explanations, it is often important to determine the exact identity of persons or objects. Many cases concern not specifically *what happened* but rather *who did it* (De Poot et al., 2004). In the case of Haaknat, if we decide to accept Haaknat's version of the events $S_h$, the question is still who robbed the supermarket. In the case study in Chapter 6, the most important question is about the identity of the killer. We can also inquire as to the identity of objects, that is, *with what* the crime was committed (again see De Poot et al., 2004)? For example, in the Rijkbloem case presented by Wagenaar and colleagues (1993) and adapted by Bex and colleagues (2007a), an important issue is the exact type of gun used and the case study also contains a discussion on the murder weapon. Evidential data sometimes provides only part of the puzzle as to the identity of some entity. For example, we might know that *someone* robbed the supermarket, or that the perpetrator was male and that he drove a red car. The exact identity must then be established by combining evidence and here evidence can be interpreted in multiple ways as to support one identity or another.

Reasoning about partially known or unknown entities is not uncommon when reasoning about criminal cases. One way to logically model this would be to include some sort of identity assumptions; for example, as part of our hypothesis we might assume that robber(supermarket) = Haaknat. Keppens and Schafer (2005) model this type of reasoning with *pegs*, expressions that refers to a unique entity whose identity can be unknown or only partially known. The current theory takes a different approach to reasoning about the identity of entities in our hypotheses, namely one that involves reasoning with abstraction generalizations.

Following Kautz (1991), Console and Dupré (1994) incorporate what they call abstraction axioms in the theory $T$ (recall that in these traditional models the hypotheticals together with the theory explain the observations, $H \cup T \vdash O$). In Console and Dupré's model, abstraction axioms, which express ISA relationships, are syntactically the same as the causal rules with material implication and can be incorporated in $T$ to explain an observation. However, they are not causal rules and should not be seen as such. Abstraction axioms allow for reasoning with models that are expressed at different levels of abstraction. In the current book, such abstraction axioms are called *abstraction generalizations*. They can be used to connect a story to a story scheme but also to connect a single event in a story to a more abstract version of the event. In this way, abstraction generalizations can be used to reason about the identity of entities. As an example, consider the situation from Fig. 4.1 (p. 86), in which the abstraction generalization Haaknat robs supermarket $\Rightarrow_A$ supermarket is robbed is used to provide the identity of the person who robbed the supermarket. Say that the supermarket is robbed is an explanandum; the combination of $S_p$ and Haaknat robs supermarket $\Rightarrow_A$ supermarket is robbed explains this explanandum whilst the story [$S_h$, Haaknat robs supermarket $\Rightarrow_A$ supermarket is robbed] does not. It is also possible that someone else robbed the supermarket and other explanations for the explanandum supermarket is robbed should also be considered. For example, there may be some evidence that another man called Johnny robbed the supermarket. With the abstraction generalization Johnny robs supermarket $\Rightarrow_A$ supermarket is robbed an alternative explanation for the explanandum supermarket is robbed can be given.

Note that in Fig. 5.4, abstraction links are rendered as gray arrows which are somewhat thicker than the causal arrows. This way of modelling reasoning about identities is natural and relatively simple compared to the formal methods described above and hence better fits the current sense-making context. In the case study in

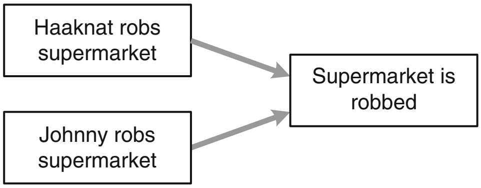



**Fig. 5.4** Two alternative explanations that use abstraction generalizations

Chapter 6, more examples of providing alternative identities of a person or an object through abstraction relations will be given.

To summarize, the current causal theory allows for the construction of hypothetical stories that explain the explananda. Similar to the way in which particular arguments based on an evidential theory *ET* can be part of an evidential argumentation theory *AT* (Definition 5.2.7), particular explanations based on a causal theory *CT* can be part of a *causal explanation theory*:

**Definition 5.3.5 (Causal Explanation Theory)**

$XT = (CT, Expl, Schemes)$ *is a* causal explanation theory, *where*

– $CT = G_{CA} \cup F \cup H$ *is a causal theory;*
– $Expl$ ($\subseteq Stories_{CT}$) *is a set of* explanations *for the explananda F;*
– *Schemes is a set of* story schemes.

Given a specific *CT*, a causal explanation theory contains explanations for the explananda in that *CT*. A causal explanation theory also contains *story schemes*, which will be discussed in the next section. As with all such definitions, the rest of this chapter always assumes some arbitrary but fixed causal explanation theory unless specified otherwise.

### 5.3.4 *Story Schemes*

An important part of holistic reasoning with stories is reasoning with *story schemes*: when reasoning with stories, the world knowledge that is used often does not have the form of individual causal generalizations but is more naturally thought of as a story scheme. These story schemes are abstract renditions of general stories about how things happen in the world. In the causal theory, story schemes can be represented using causal generalizations from $G_{CA}$ and particular stories can then be connected to a scheme with abstraction generalizations. In Section 3.2.2, story schemes and the ways in which stories can correspond to such schemes were informally discussed. This section contains the formal definitions of the correspondence between stories and schemes. The criteria that a good story should be *complete*, that is, that it should correctly match a story scheme will be further discussed below in Section 5.4.3.

A story scheme can be defined as a collection of literal schemes and (causal) generalization schemes:

**Definition 5.3.6 (Story Scheme and its Components)**

– *A story scheme $G_S \in$ Schemes is a set comprised of literal schemes and causal generalizations.*
– *For any scheme $G_S$, the set of Components*$(G_S) = G_S \cup \{\varphi$ is a literal scheme | $\varphi \in G_S$ or $\varphi$ is the antecedent or the consequent of some $g_i \in G_S\}$.

The individual components of a scheme are similar to the events in a story, that is, all the possible individual literal (schemes). Recall from Definition 5.1.1 that a generalization with (free) variables is a scheme for all its ground instances. Analogously, a literal scheme is a scheme for all its ground instances; for example, *x* robs *y* is a scheme for Haaknat robs supermarket and John robs bank. Now, a story scheme can contain only literal schemes, for example, the robbery scheme from Fig. 3.22 (p. 68): {*x* wants *g*, *y* has *g*, *x* wants to rob *y*, *x* has the opportunity to rob *y*, x robs *y*, *y* loses *g*}. In this way, story schemes of various degrees of specificity can be represented. For example, the story scheme for bank robberies is a variant of the above scheme where *y* substituted for bank. Story schemes can also contain predicates of arity 0, for example {motive, goal, action, consequence}. This intentional action scheme was already given in Fig. 3.20 on p. 64 with explicit causal links and a formal story scheme can also contain causal links: {motive $\Rightarrow_C$ goal, goal $\Rightarrow_C$ action, action $\Rightarrow_C$ consequence}. Notice that in this last scheme, the individual literals motive, goal, action and consequence are also components.

Events in a story can be linked to the components of a scheme by instantiating elements of a literal scheme or by using abstraction generalizations; the links between events and components of a scheme are called *correspondence* links:

**Definition 5.3.7 (Correspondence)**

*A set of events E* corresponds to *a literal scheme $\varphi$ and vice versa iff for a ground instance of $\varphi$: $E \cup G_A \vdash_A \varphi$.*

So a set of events corresponds to a literal scheme (i.e. a component of a story scheme) if a ground instance of the scheme is $\mathcal{A}$-derivable (Definition 5.3.2) from the events. In some cases, an explicit abstraction generalization is not necessary. For example, the event Haaknat robs supermarket is itself a ground instance of *x* robs *y* so it corresponds to this particular component of the robbery scheme. In other cases, an explicit abstraction generalization is needed. For example, Haaknat robs supermarket corresponds to the action component of the intentional action scheme through the abstraction generalization Haaknat robs supermarket $\Rightarrow_A$ action because {Haaknat robs supermarket} $\cup$ {Haaknat robs supermarket $\Rightarrow_A$ action} $\vdash_A$ action. Multiple events can correspond to one component of the scheme. For example, the events {Haaknat is a drug addict, Haaknat needs money} correspond to the motive component of the intentional action scheme through the generalization Haaknat is a drug addict $\wedge$ Haaknat needs money $\Rightarrow_A$ motive. Furthermore, the derivation from the event to the component of the scheme can be more than one step. In this way, a story can correspond to multiple story schemes and story schemes can correspond to each other. For example, the event Haaknat is at the supermarket corresponds to components of the robbery scheme and the intentional action scheme as given in Fig. 3.22 (p. 68) through the following derivation: [Haaknat is at the supermarket, Haaknat is at the supermarket $\Rightarrow_A$ $x$ has the opportunity to rob $y$, $x$ has the opportunity to rob $y \Rightarrow_A$ physical state].

The above Definition 5.3.7 only defines correspondence between some events and some components of a story scheme and does not say how a story corresponds to a generalization in a story scheme. The correspondence of a story to a scheme often depends not only on the correspondence of the individual events to some component of the scheme but also on the correspondence between the causal structure of the scheme and the story. With story schemes which have no explicit causal structure (i.e. story schemes which do not contain generalizations) this is not an issue. However, when a story scheme does have such a structure the corresponding story should conform to this. Take, for example, the following story scheme: {$x$ commits crime $\Rightarrow_C$ $x$ hides, $x$ hides $\Rightarrow_C$ police find $x$}. This "commit-crime-and-hide-scheme" expresses the knowledge that it is not unusual for people who commit crimes to hide but that the police will usually find them. In this scheme, explicit causal information is expressed through the generalizations: committing the crime leads to hiding which in turn leads to being found. Now, say that the event Haaknat robs supermarket corresponds to $x$ commits crime, the event Haaknat hides in park corresponds to hides and the event police find Haaknat in park corresponds to police find $x$. Because of the causal structure of the scheme, for the story $S_p$ about Haaknat to correctly match the scheme it would in this case also be required that somehow Haaknat robs supermarket leads to Haaknat hides in park which in turn leads to police find Haaknat in park. In many cases a causal generalization in a story scheme will not be directly recreated in a story. For example, there is no generalization Haaknat robs supermarket $\Rightarrow_C$ Haaknat hides in park in the Haaknat story $S_p$ that in some way directly corresponds to commit crime $\Rightarrow_C$ hide in the scheme. However, the story corresponds to the scheme because there is a *chain* of causal relations in the story that corresponds to the causal generalization commit crime $\Rightarrow_C$ hide. This type of correspondence between a generalization in a story scheme and a story is defined as follows:

**Definition 5.3.8 (Correspondence of Generalizations)**

*A story S* corresponds to *a generalization g*: $\varphi \Rightarrow_C \psi$ *and vice versa iff*

(1) *there is a set of events* $E_1 \subseteq Events(S)$ *such that* $E_1$ *corresponds to* $\varphi$*; and*
(2) *there is a set of events* $E_2 \subseteq Events(S)$ *such that* $E_2$ *corresponds to* $\psi$*; and*
(3) *there is a substory S' of S such that:*
    (a) $E_1 \subseteq IniEv(S')$*; and*
    (b) $E_2 \cap IniEv(S') = \emptyset$*; and*
    (c) *For all* $e \in E_2$: $(IniEv(S') \setminus E_1) \cup Gens(S) \nvdash_{CA} e$.

In words, a story $S$ corresponds to a generalization $g$ and vice versa if and only if both the antecedent and the consequent of $g$ correspond to a set of events from $S$ (conditions 1 and 2) and the set of events $E_2$ that corresponds to the consequent of $g$ directly follow from the set of events $E_1$ that corresponds to the antecedent of $g$. The requirement that the events in $E_2$ directly follow from $E_1$ is ensured by condition 3 as follows. First, in order to ensure that $E_2$ does not follow from earlier events than those in $E_1$, take the substory $S$ such that the events in $E_1$ are initial events (condition 3a); this does not have to be a proper substory, so it is possible that $S' = S$. The events in $E_2$ should not be initial events in $S'$ so that the events in $E_2$ cannot be derived from $E_2$ itself (condition 3b). Condition 3c then ensures that events $E_1$ are needed to derive all the events in $E_2$, as without $E_1$ one or more events in $E_2$ are not derivable given the story $S$.

Let us return to the example of the story $S_p$ and the story scheme {*x* commits crime $\Rightarrow_C$ *x* hides, *x* hides $\Rightarrow_C$ police find *x*}. Above it was already argued that events in the story correspond to components of the story scheme. The question here is whether the story also conforms to the causal structure as set by the scheme. Consider Fig. 5.5, in which the correspondence between part of the story and the story scheme is represented. As can be seen, in this case all the conditions are met for the story to correspond to the generalizations *x* commits crime $\Rightarrow_C$ *x* hides and *x* hides $\Rightarrow_C$ police find *x* in the scheme, that is, for the story to conform to the scheme's causal structure. The story corresponds to generalization *x* commits crime $\Rightarrow_C$ *x* hides as follows: first, Haaknat robs supermarket corresponds to *x* commits crime (condition 1) and Haaknat hides in park corresponds to *x* hides (condition 2). Now, the substory Haaknat robs supermarket $\wedge$ police arrive $\Rightarrow_C$ Haaknat gets into car and takes off, Haaknat gets into car and takes off $\wedge$ Haaknat thinks police are following $\Rightarrow_C$ Haaknat hides in park has Haaknat robs supermarket in its initial events (condition 3a), does not have Haaknat hides in park in its initial events (condition 3b) and without Haaknat robs supermarket the event that corresponds to the generalizations consequent, Haaknat hides in park, cannot be derived. For the other generalization in the scheme, *x* hides $\Rightarrow_C$ police find *x*, it is easier to see that all conditions from Definition 5.3.8 are met because the causal generalization Haaknat hides in park $\wedge$ police search in park $\Rightarrow_C$ police find Haaknat in park almost directly matches the generalization in the scheme *x* hides $\Rightarrow_C$ police find *x*, save for the fact that the generalization in the story contains an additional literal in its antecedent.

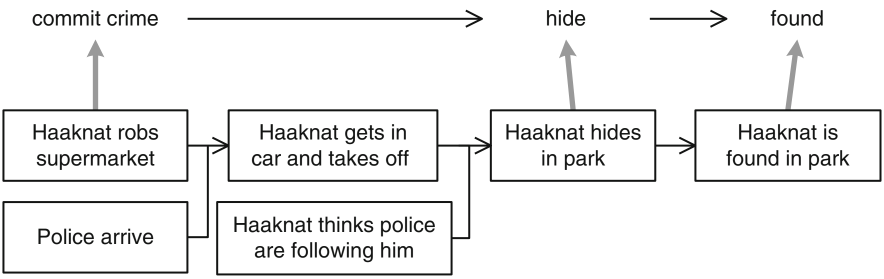



**Fig. 5.5** Graphical representation of matching the Haaknat story to the "crime-and-hide" story scheme

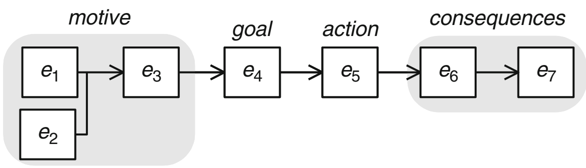



**Fig. 5.6** Graphical representation of story schemes and stories

In Fig. 5.5, the way in which a story is matched to a scheme is rendered graphically. In Here, the elements of the scheme are represented as text and the abstraction links as thick grey arrows. In the rest of this book, the abstraction links linking a story to a scheme will sometimes not be explicitly rendered and if more events correspond to one part of a scheme, this can be visualized with a grey rounded box, viz. Fig. 5.6. In this way, complex stories and story schemes can be visualized clearly.

The various ways in which a story and a scheme can correspond have now been formally defined. In the hybrid theory, story schemes can be used for different purposes. For example, they can help in the construction of stories by serving as a template; examples of this will be given in Section 5.6 and in Chapter 6. Another important use of schemes is their use in determining the coherence (and thus the quality) of stories. This will be further discussed and defined in Section 5.4.3.

## 5.4 A Hybrid Theory of Argumentation and Explanation

In the previous sections, the evidential (argumentative) and the causal (story-based) parts of the hybrid theory have been defined. The argumentative part consists of input data *ET* (evidential data and evidential generalizations) from which arguments can be constructed. These arguments can attack each other; the combination of arguments based on an *ET* and their attack relations is an evidential argumentation theory *AT*. The arguments in an *AT* can be assigned a status overruled, defensible or justified according to some argument-based semantics. The causal part *CT* consists of input data *H* and $G_{CA}$ (hypotheticals, causal generalizations and abstraction generalizations) from which hypothetical causal stories that explain the explananda *F* can be constructed. Abstraction generalizations allow for reasoning about stories on various levels of specificity and also allow a story to be connected to a story scheme; this connection between a story and a scheme is known as correspondence. The combination of explanations for the explananda and possible story schemes is called a causal explanation theory *XT*.

Now that both the argumentative and the story-based part of the hybrid theory have been defined separately, the combination of these theories, which was informally discussed in Chapter 4, can be given. The hybrid theory is a combination of an evidential argumentation theory and a causal explanation theory:

**Definition 5.4.1 (Argumentative Explanation Theory)**

*A hybrid argumentative explanation theory is a tuple* $AET = (AT, XT)$*, where:*

– $AT = (ET, Args, Attack)$ *is an evidential argumentation theory;*
– $XT = (CT, Expl, Schemes)$ *is a* causal explanation theory*, where*
– $CT = G_{CA} \cup F \cup H$ *is a causal theory such that every* $f \in F$ *is the conclusion of a justified argument* $A$ *in* $Args$ *and* $A$*'s premises contain at least one* $\varphi \in Evidence(A)$.

An argumentative explanation theory *AET* is essentially the information which allows us to construct and compare stories, arguments and their combination. In the *AET*, the set $I_E \subseteq ET$ contains the evidence, the necessary facts which cannot be called into question. The combination of $G_E \subseteq ET$, $G_{CA}$ and $H$ (both subsets of *CT*) contains the contingent facts, facts which are assumed but can be attacked and defeated. The combination of all these contingent fact essentially represents the current *stock of knowledge* about which a cognitive consensus should be reached. The rest of the definitions in this chapter assume some arbitrary but fixed hybrid theory *AET*, unless explicitly mentioned. This theory has as its basic language the defeasible language $\mathcal{L}_\mathrm{d}$ and uses the logic $\mathcal{D}_\mathrm{E}$ (the defeasible logic extended with the evidential reasons). The argumentative part *AT* is as defined in Definition 5.2.7 and the story-based part in Definition 5.3.5.

In the hybrid theory *AET*, the argumentative part directly influences the composition of the story-based part of the theory. This is evident in the above definition, where the explananda should follow from evidential data through a justified argument. The reason for this is that in police investigation we are only interested in states or events which have actually happened. Therefore, events which do not follow from evidence (i.e. we are not sure they happened) do not have to be explained (see p. 92).

Arguments from the *AT* also play and important role in the comparison of explanations in *XT* according to their coherence and the extent to which they conform to the evidence. Arguments can be used to support or contradict an explanation with evidence, increasing the explanation's evidential support or contradiction, respectively. Arguments based on the stock of knowledge can also be given for or against elements in a story, increasing or decreasing the story's causal plausibility. In the following sections, the criteria for determining the quality of a story, which were presented informally in Sections 4.2 and 4.3, will be formally defined. First, the extent to which a story conforms to the evidence will be discussed in Section 5.4.1; Section 5.4.2 discusses a story's plausibility and in Section 5.4.3 the criterion that a story should conform to some plausible story scheme will be defined.

## *5.4.1 Supporting and Contradicting Stories*

In Section 4.2, three criteria for determining the extent to which a story conforms to the evidence were given, namely *evidential support*, *evidential contradiction* and *evidential gaps*. Evidential support concerns the pieces of evidence that support some event or causal relation in a story, evidential contradiction concerns the pieces of evidence that contradict some event or causal relation in a story and the evidential gaps are the events of a story which are unsupported and not contradicted by evidence.

In order to determine evidential support, a notion of direct support is needed:

**Definition 5.4.2 (Direct Support)**

*$\varphi$ (justifiably or defensibly)* supports *$\psi \neq \varphi$ iff there is a (justified or defensible) argument $A \in Args$ such that $\varphi \in Prem(A)$ and $\psi \in Conc(A)$ and $Prem(A) \setminus \{\varphi\}$ $\nvdash_{\mathsf{E}} \psi$.*

In words, a premise of a non-overruled argument supports its conclusions, provided that the premise is needed in the argument to infer the conclusion. For example, if argument (*H in car*) (Section 5.2.2) is not overruled, then we can say that $e_1$ supports Haaknat got into the car. That the conclusion is inferred from the premise is enforced by the condition that *without* the particular premise, the conclusion is not derivable from the premises of the argument. For example, the definition of arguments (Definition 5.2.4) allows us to construct a new argument by adding a premise, say a piece of evidence $e_2$, as a new line of argument to (*H in car*). However, without this evidence it would also be possible to infer that Haaknat got into the car so it has no influence on the derivation of Haaknat got into the car and it should not count as supporting it. The definition of support allows us to determine which conclusions are *directly supported* by the elements from the input of the evidential argumentation theory *AT*. There are 2 degrees of support, justified and defensible support, where justified is stronger than defensible.

The definition of support allows us to say, for example, that a piece of evidence supports an event in a story. The evidential support can now be defined as the set of all pieces of evidence that support some event or causal relation in a story.

**Definition 5.4.3 (Evidential Support)**

*The* evidential support *of a story S is the set* $E^{+}(S) = \{\varphi \in I_E \mid \varphi$ supports some $\psi \in S\}$.

Here, justified and defensible evidential support can be distinguished, which correspond to the justified or defensible support that the evidential data gives to the element in the story. For the moment, however, this distinction will not be made and any piece of evidence that supports the story through a non-overruled argument simply supports the story. As an example, say that we want to explain the explanandum police find Haaknat in park and the set of explanations is $\{S_p, S_h\}$ (see p. 134 and 136). Furthermore, assume the set of evidence $I_E$ to be $\{e_{\mathrm{w1}}, e_{\mathrm{H}}, e_{\mathrm{pr}}\}$, where $e_{\mathrm{w1}}$ is a witness testimony about a Haaknat look-alike entering the car, $e_{\mathrm{H}}$ is Haaknat's denial about entering the car and $e_{\mathrm{pr}}$ is a police report detailing the police's response to the robbery and the fact that Haaknat was found in the moat. Finally, the set of arguments *Args* is $\{$(*H in car*), (*H not in car*), $A_{\mathrm{pr}}\}$, where (*H in car*) and (*H not in car*) are as on p 123 and 126 and $A_{\mathrm{pr}}$ is a simple argument with

$e_{pr}$ as its premise and police arrive, police search in park, police find Haaknat in park as its conclusions. The evidential support of $S_p$ is now $\{e_{w1}, e_{pr}\}$, because (*H in car*) is defensible and $A_{pr}$ is justified. Explanation $S_h$ is currently only supported by $e_{pr}$. However, recall that it was Haaknat's own testimony on which $S_h$ was based. A new piece of evidence $e_{HS}$ can be added which is Haaknat's testimony in which he gives an account of the episode with Bennie and his flight. This evidence then supports all the events in $S_h$ through a simple argument based on the witness testimony reason.[13] The evidential support of $S_h$ is then $\{e_{HS}, e_{pr}\}$.

Evidential contradiction is essentially the opposite of support: all the pieces of evidence that contradict some event or causal relation in a story. Again, a general notion of contradiction is needed; however, due to the nature of "undercutting contradiction" (where DMP on a causal generalization is undercut), the definition of contradiction explicitly assumes a story:

**Definition 5.4.4 (Contradiction)**

*Given a story S,* $\varphi$ (*justifiably or defensibly*) contradicts $\psi \in S$ *iff*

- $\varphi$ (*justifiably or defensibly*) *supports* $\overline{\psi}$*; or*
- $\varphi$ (*justifiably or defensibly*) *supports* $\ulcorner\neg(\{\chi_1,\ldots,\chi_n\} \gg \psi)\urcorner$ *and* $\psi$ *is obtained from some earlier events* $\chi_1,\ldots,\chi_n$ *in S by the application of a prima facie reason.*

So an argument contradicts a proposition if it supports the complement of that proposition, or if it undercuts a causal relation in the story, that is, if it is an argument for an exception or the invalidity of a causal generalization used in the story. Like support, contradiction has 2 degrees, justifiable and defensible, and a premise of an overruled argument does not contradict its conclusion. The notion of contradiction is very similar to the notion of attack from Definition 5.2.6. However, attack is defined between two lines of argument and contradiction is defined between an argument $A$ and an element in a story $S$. Evidential contradiction is now the set of all pieces of evidence that contradict some element of a story (Definition 5.4.5).

**Definition 5.4.5 (Evidential Contradiction)**

*The* evidential contradiction *of a story S is the set* $E^-(S) = \{\varphi \in I_E \mid \varphi$ contradicts some $\psi \in S\}$.

This definition is essentially the opposite of evidential support. Where a story is better if it covers more of the evidence, it is worse if it is contradicted by more evidence. Similar to the definition of evidential support, justified and defensible evidential contradiction can be distinguished but again this distinction will not be made. In the above example, the piece of evidence that directly contradicts the explanation $S_p$ is $e_H$. $S_h$ is not directly contradicted.

Notice that only evidence that *directly* supports or contradicts a story is taken into account when considering evidential support or contradiction. For example,

[13] Recall that a suspect's testimony is here also regarded as a witness testimony.

say that (*H not in car*) is undercut by some evidential data $e_{\mathrm{Hv}}$ against Haaknat's veracity. This $e_{\mathrm{Hv}}$ does not count towards $S_p$'s evidential support even though it is compatible with the story and in a sense "defends" the story by defeating one of the contradicting arguments. Here the story is only *indirectly supported* by this evidence. Similarly, even though the arguments (*common appearance*) and (*invalidity* $g_{\mathrm{L}}$) (Section 5.2.3) can be used to undercut (*H in car*) and thus decrease $S_p$'s support, the evidence on which they are based ($e_{\mathrm{ca}}$ or $e_{\mathrm{e2}}$) does not count towards $S_p$'s evidential contradiction, because it is only indirectly incompatible with the story.

The third criterion that pertains to a story's conformance to the evidence is that of *evidential gaps*. Recall that evidential gaps are those events for which there is no evidence from which we may infer either that the event happened or did not happen. That is, the evidential gaps are those states and events which are not directly supported or contradicted by evidential data:

**Definition 5.4.6 (Evidential Gaps)**
*The* evidential gaps *of a story S are the elements of a set* $E^G(S) =$
$\{a \in Events(S) \mid \neg\exists\varphi \in I_E \text{ such that } \varphi \text{ supports or contradicts } a\}$

Note how the evidential gaps is a set of events rather than a set of evidential data and that causal generalizations that are unsupported do not count towards the evidential gaps of a story. If we do not take (*common appearance*) or (*invalidity* $g_{\mathrm{L}}$) into account, then all elements in the example story $S_p$ are gaps except for Haaknat got into the car, police arrive, police search in park, police find Haaknat in park. It does not matter if either (*H in car*) or (*H not in car*) is somehow overruled. As long as one of these arguments is not overruled, the event Haaknat got into the car has evidence that is of importance to it. The explanation $S_h$ contains no evidential gaps because all events are supported by either $e_{\mathrm{HS}}$ or $e_{\mathrm{pr}}$. If, however, the inference from Haaknat's testimony is somehow undercut, all events are gaps except for those that follow from $e_{\mathrm{pr}}$.

### *5.4.2 The Coherence of Stories: Plausibility and Implausibility*

In Section 4.3 three criteria for the coherence of a story were given, *internal consistency*, *plausibility* and *completeness*. The criterion of internal consistency was already incorporated as condition 2 in the definition of explanations (Definition 5.3.4). Recall that with this criterion, it is not the extent to which a story is consistent but rather the question *if* a story is consistent: explanations that internally contradict themselves can never be fully supported because this would require us to accept contradictory evidence. Therefore, we model consistency as a condition on explanations so that explanations which are inconsistent are not considered. Plausibility and implausibility are defined in this section and completeness will be discussed in Section 5.4.3.

The *plausibility* of a story is the extent to which the causal generalizations and evidential gaps are supported by explicit arguments based on the stock of knowledge

$G_E$. Arguments based on $G_E$ can support as well as contradict an element of a story, so it is possible to define the plausibility as well as the implausibility of a generalization:

**Definition 5.4.7 (Plausibility)**

*The* plausibility *of a story S is a set* $P^+(S) = \{\varphi \in S \mid \varphi$ is supported by some $\psi \in G_E$ and there is no $\chi \in I_E$ such that $\chi$ supports $\varphi\}$.

Notice that plausibility is only relevant for causal relations and events not supported by evidence. This is because plausibility is a notion which is essentially established independently of the evidential data. In an ideal situation, evidential data is provided for the elements of a story. If there is, for example, evidential data that supports an event, we do not need to reason from our stock of knowledge in order to make this particular event more plausible. However, when there is no evidential data supporting an event it is important to consider whether the event is plausible, that is, if it can be inferred from our general stock of world-knowledge. In Fig. 5.7, the generalization $g_{c2}$ is plausible (and hence part of the "plausibility set") because it is supported by the general knowledge Haaknat is the kind of man who acts on his impulses. The expert statistical evidence in this figure does not count towards the story's plausibility (it does, however, count towards the story's evidential support).

Just as evidential contradiction is the inverse of evidential support, *implausibility* is the inverse of plausibility:

**Definition 5.4.8 (Implausibility)**

*The* implausibility *of a story S is a set* $P^-(S) = \{\varphi \in S \mid \varphi$ is contradicted by some $\psi \in G_E$ and there is no $\chi \in I_E$ such that $\chi$ contradicts $\varphi\}$.

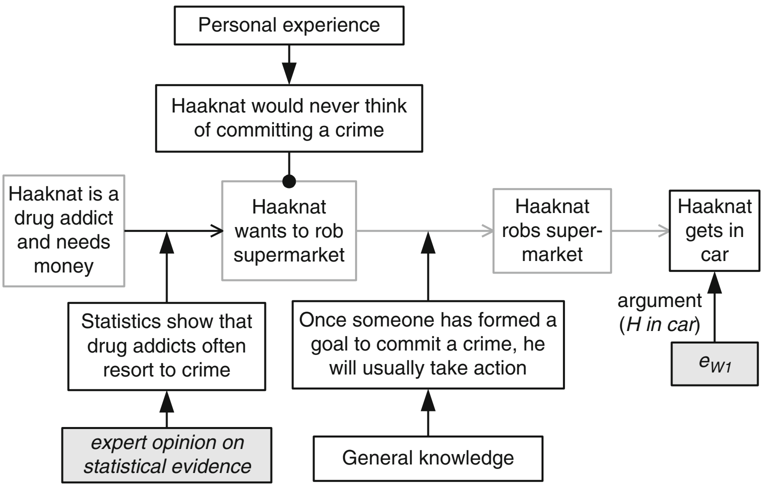



**Fig. 5.7** Graphical representation of supporting and contradicting a story

Implausibility is only relevant for elements of a story that are not already contradicted by evidential data. The reason for this is similar as with plausibility: if there is evidential data that directly contradicts an event, we do not need to reason from general knowledge in order to make this particular event more implausible. As an example, consider Fig. 5.7; here Haaknat wants to rob supermarket is implausible because it is contradicted by a piece of knowledge from personal experience. The combination of stories and their supporting arguments can be visualized. Take, for example, a part of the prosecution's Haaknat story $S_p$ and its supporting arguments (Fig. 5.7).

In Fig. 5.7, part of the prosecution's story and supporting and contradicting arguments are shown. The evidence $e_{W1}$ supports Haaknat gets in car through argument (*H in car*). The expert statistical evidence leads to the conclusion valid($g_{c1}$), where $g_{c1}$ is (Haaknat is a drug addict $\wedge$ Haaknat needs money $\Rightarrow_C$ Haaknat wants to rob supermarket); this has been rendered as the argument based on the expert evidence directly supporting the generalization $g_{c1}$. Similarly, The general knowledge that Once somebody has formed a goal to commit a crime, he will usually take action supports the generalization $g_{c2}$: Haaknat wants to rob supermarket $\Rightarrow_C$ Haaknat robs supermarket. Haaknat would never think of committing a crime contradicts the event Haaknat wants to rob supermarket. Notice that the event Haaknat gets in car and the causal link expressed by generalization $g_{c1}$ are rendered with a black line because they are directly supported by evidence. Evidential gaps and unsupported causal links, which are not supported by evidence, have a gray line ($g_{c2}$ is supported by an argument but not by evidence). In this way, they are distinguished from events or links supported by evidence.

### 5.4.3 *The Coherence of Stories: Story Schemes and Completeness*

Given a story and a scheme, the notion of completeness has essentially two sides: a story *completes* a story scheme if for every element of the scheme it has a corresponding state or event, and a story *fits* a story scheme if all elements of the story correspond to some element in the scheme. In other words, a story is complete if it "has all its parts" and fits if it has no "loose ends". The correspondence between stories and generalizations or individual components of a scheme was defined in Section 5.3.4. In this section, it will first be defined how a story completes and fits a scheme and then the criterion of completeness will be defined. A story *completes* a story scheme if and only if for every element of the scheme it has a corresponding state or event:

**Definition 5.4.9 (Completing a Scheme)**

*A story $S \in Expl$* completes *a story scheme $G_S \in Schemes$ iff*

– *for each individual component $\varphi \in Components(G_S)$ there is a set of events $E \subseteq Events(S)$ such that $E$ corresponds to $\varphi$; and*
– *each generalization $g_i \in G_S$ corresponds to $S$.*

In the example in Fig. 5.5 (p. 131), the story $S_p$ completes the crime-and-hide-scheme. Note that the second condition (that each generalization in the scheme corresponds to the story) ensures that the story has the same causal structure as the scheme.

A story *fits* a story scheme if and only if all elements of the story correspond to some element in the scheme:

**Definition 5.4.10 (Fitting a Scheme)**

*A story $S \in Expl$* fits *a story scheme $G_S \in Schemes$ iff for all* e $\in Events(S)$, *there is some set $E$ (*E $\in E$) *such that $E$ corresponds to some $\varphi \in Components(G_S)$.*

Notice that the definition of fitting does not enforce that the scheme has the same causal structure as the story but merely that the story has no loose events which do not correspond to an element of a scheme. For example, take the following story: [Johnny hides, Johnny hides $\Rightarrow_C$ police find Johnny, police find Johnny $\Rightarrow_C$ Johnny kills the officer who finds him]. This story can be said to fit the "commit-crime-and-hide-scheme", because Johnny hides corresponds to *x* hides, police find Johnny corresponds to police find *x* and Johnny kills the officer who finds him corresponds to *x* commits crime. However, the schemes causal structure is not the same as that of the story: in the story, Johnny commits the crime after he hides (because he is found) and in the scheme, the crime is committed before the culprit hides. This is not a problem, because completing and fitting will be used together to define the criterion of completeness below.

One requirement that is important in the notion of completeness is that a story should not complete and fit just any story scheme but that it should complete and fit a *plausible* story scheme. The story schemes that are part of the set *Schemes* are not automatically plausible: it is possible to reason about their plausibility with arguments in the same way as the plausibility of stories can be reasoned about. In this way, a cognitive consensus on the story schemes can be reached. Such arguments will be mainly based on the stock of knowledge. For example, a restaurant story scheme that contains the generalization $g_r$: *x* eats in restaurant $\Rightarrow_C$ *x* leaves without paying can be attacked by the following argument:

*Argument* (*restaurant*)

(1) $g_r$: $\Rightarrow_E$ people do not leave restaurants without paying. (***$G_E$***)
(2) $g_{e3}$: people do not leave restaurants without paying $\Rightarrow_E$ $\neg valid(g_r)$ (***$G_E$***)
(3) $\neg valid\ (g_r)$. (***1, conclusive***)

Schemes which are contradicted in this way should not be used to test a story's completeness. Accordingly, the set of *plausible schemes* should be defined:

**Definition 5.4.11 (Plausible Story Scheme)**

*A scheme $G_S \in Schemes$ is a* plausible scheme *iff there is no $\chi \in ET$ that contradicts a ground instance of a generalization in $G_S$ or a component of $G_S$.*

Note that a scheme can be contradicted by both evidence and other information from input and that an overruled argument does not contradict a scheme. The formal criterion that a good story should be *complete* can now be defined as follows:

**Definition 5.4.12 (Completeness)**

*A story $S \in Expl$ is* complete *iff it completes and fits a plausible story scheme $G_S \in Schemes$.*

Note that this definition depends on the plausible schemes that are part of the hybrid theory *AET*. The problem that fitting did not require the causal structure of the story and the scheme to be similar is now obviated by the fact that completing requires the causal structures to be similar and completeness requires both completing and fitting. The example prosecution story about Haaknat $S_p$ completes the crime-and-hide-scheme but does not fit this scheme: there are a number of loose ends which do not fit the scheme. For example, why did Haaknat need the money? In order for the story to fit the scheme, the scheme would have to be extended so that such events (i.e. Haaknat needs money) correspond to an element of the scheme. Examples of complete stories will be given in the example of the dialogue game in Section 5.6 and in the case-study in Chapter 6.

## *5.4.4 Assessing and Comparing Stories*

In the above sections, the hybrid theory has been formally defined. Arguments that can be used to reason with and about evidential data and the stock of knowledge and generalizations were discussed in Section 5.2; stories and the way in which they explain the explananda were covered in Section 5.3. In Section 5.4 the combination of arguments and stories was defined and the criteria by which the quality of stories can be judged were formally defined. With the formal and precise definitions of the criteria it is now possible to precisely determine whether and to what extent a story conforms to the evidence and whether a story is coherent. In other words, the formal definitions of the criteria allow the critical questions from Section 4.4 to be answered in detail.

In Section 4.4, various ways of comparing stories using the criteria were discussed. One such way is to determine the extent to which the stories conform to the evidence and the extent to which a story is coherent and then provide an ordering on stories. However, as was argued before, there are difficulties in comparing stories in such a discreet and mathematical way and it is better to use such orderings to *guide* a process in which an analysis of the evidence and the stories in a case is performed. The dynamic part of the hybrid theory, which determines how the criteria can guide such an analysis and how an analysis can be developed and refined, will be discussed in the next section.

## 5.5 Dialogues About Proof

The hybrid theory as proposed in the previous sections discusses a *static* viewpoint: the current status of an argumentative and explanation-based analysis of a case can be determined but the dynamics of developing and refining such an analysis have not yet been discussed. The analysis of stories and evidence in a case is a process and exactly how this process of proof takes form depends on the persons that perform the analysis and the specific legal context the analysis is performed in. In a decision-making context, for example, the defence is confronted with a complete story about what happened, namely the prosecution's story. Usually, this story is already supported by evidence and the defence will try to attack the prosecutor's evidential arguments (by arguing, for example, that a witness is not trustworthy) or the defence gives an alternative explanation (for example, that an alleged killer acted in self-defence). In an investigation context, however, things are different. Often, a team of criminal investigators is faced with some initial evidence and they construct several possible stories (or scenarios) and then try to find new evidence that supports or discredits these scenarios. During the investigation there is constant interaction between the scenarios and the evidence: a scenario provides a frame in which new evidence can be interpreted and, at the same time, new evidence is used to support or discredit a scenario or to extend a scenario.

The dynamics of the process of proof can be modelled as a *dialogue game*, which provides a set of rules or principles for a coherent dialogue. There are various types of dialogue games. For example, argumentative dialogue games for persuasion (see Section 3.1.2, see Prakken, 2006 for an overview), where one of the players makes a claim which he has to defend while the other player's goal is to dispute this claim, or dialogue games for negotiation (see Rahwan et al., 2004), where the goal is to negotiate about some "good". As was already discussed, in a dialogue game for the process of proof is an instance of an *inquiry dialogue*. The important feature of the proposed dialogue game is that the players jointly build a hybrid theory *AET*; thus the goal of an inquiry dialogue – namely collecting, organizing and assessing hypothetical stories and evidence – is collectively pursued by the players. In the dialogue game, the players should be able to both *explain* why or how something happened and *argue* that one account of what happened should be accepted given the current evidence and general knowledge. In an inquiry dialogue, none of the players really wants to "win", since the goal of the dialogue is to build an *AET* and the players all want to find the best explanation for the explananda. However, disregarding the adversarial part of the dialogue increases the risk of tunnel-vision. The adversarial setting can be enforced by demanding that the players constantly aim to have an explanation which is better than the other players' explanations. In order to achieve this, they will then extend and support their own explanation and attack the other players' explanations.

The aim of this section is to model the dynamics of the process of proof as a *formal dialogue game*. In this dialogue game, it should be possible to build, analyse and change stories and their supporting arguments. In addition to describing the dynamic process of proof, the dialogue game also is a useful tool given the current notion of procedural rationality: the formal dialogue game aims to define a proper process through which beliefs about evidence, arguments and stories can be formed. In this process, an inquiry dialogue, the utterances by the players should at least be *relevant*. Furthermore, the players should be encouraged to improve their own explanations or disprove the other players' explanations. A story and its supporting evidence can then be accepted if they cannot be successfully challenged by other players in the dialogue game. Note that, unlike some dialogue games for argumentation (e.g. Loui, 1998, Prakken and Sartor, 1997), the current formal dialogue game is not intended as a formal proof theory but rather as a general framework that guides the players in a rational discussion of the evidence. Bex and Prakken (2008) have defined a formal dialogue game for inquiry dialogue based on an early variant of the hybrid theory for explanations and arguments. In this dialogue game, the basic rules for a proper discussion of evidence and stories were given. In this section, an extension of this dialogue game is presented.

### *5.5.1 Framework for a Formal Dialogue Game*

In a dialogue the players build a hybrid theory *AET* by giving explanations for the explananda, making predictions on the basis of an assumed explanation and moving arguments to support or attack explanations or other arguments. In this way, an *AET* is continually updated. Because the dialogue framework is based on a hybrid theory *AET*, the *topic language* of the dialogue game, which determines the possible content of utterances by a player, is the language $\mathcal{L}_\mathrm{d}$ and the underlying logic is $\mathcal{D}_\mathrm{E}$. The theory *AET* determines the topic of discussion: *what* is discussed during the dialogue. *How* this discussion takes form is determined by the dialogue game that is to be defined below.

In a dialogue game, the *players* build a hybrid theory *AET* by performing *speech acts* from a communication language. With these speech acts, explanations can be given for the explananda, predictions can be made on the basis of an assumed explanation and arguments can be moved to support or attack explanations or other arguments. An important part of the dialogue game is the *protocol*, which specifies the allowed moves at a certain point in the dialogue. Such a protocol is essentially a normative model for how the process of an analysis of evidence and explanations should take place: in order to be able to draw a rationally sound conclusion from an inquiry dialogue, the game should follow a specific set of rules. These rules make sure that, for example, they ensure that all explanations or arguments proposed by the players are relevant to the topic and that each of the players gets their turn in trying to improve or discredit explanations.

The dialogue game also has *commitment rules*, which specify the effects of a speech act on the propositional commitments of the players. For instance, explaining the explananda with an explanation commits the speaker to the explanation and retracting a previously moved argument removes this argument from the speaker's commitments. Commitments can be used in the protocol to constrain the allowed moves, for example, to disallow moves that make the speaker's commitments inconsistent. They can also be used to define *termination* and *outcome* of a dialogue. Because the aim of the individual players is to support and defend their own explanation, the dialogue terminates if all players are committed to the same explanation. As in any inquiry dialogue, the ultimate objective of the game is to collect, organize and assess stories and evidence. The outcome of a dialogue is therefore an *AET* which represents a collection of stories and arguments based on evidence.

For non-terminated dialogues a notion of the *current winner* can be defined; this is the adversarial element of the dialogue. The current winner is the player that is committed to the explanation that is currently the best and all players in the dialogue essentially want to become the current winner. The best explanation is determined by referring to two orderings on explanations. One of these orders explanations according to set inclusion using the criteria from Sections 5.4.1 and 5.4.2 and the other orders explanations by considering the extent to which the criteria apply as an absolute number. The notion of winning is used to control *turntaking*, with a rule that a player is to move until he has succeeded in becoming the current winner (cf. Loui, 1998). The general framework for the dialogue game can now be defined as follows:

**Definition 5.5.1 (Dialogue Game)**
*A dialogue system is a tuple DG* = (*Players*, $L_c$, *Moves*, *U, C*, <, $<_a$, *Winner*, *Turn*, *P*).

In the above definition, *Players* is the set of players, $L_c$ is the communication language, *Moves* the set of possible moves in a dialogue, *U* is an update function that determines the makeup of the theory *AET*, *C* is a function that determines the players' commitments, < and $<_a$ are orderings on explanations, *Winner* is a function that determines the current winner of the game, *Turn* is a function that determines whose turn it is to move and *P* is the protocol that determines the admissible moves. Each of these parts of the dialogue game will be defined below.

### 5.5.2 Players, Language and Moves

Most dialogue games allow for only two players. In, for example, persuasion this makes sense as there is a clear proponent and opponent arguing for and against some claim, respectively. In inquiry, however, no clearly defined positions are available at the beginning of the dialogue. Ideally, the game supports discussions between larger groups of players. In fact, the more players are in the dialogue, the more alternative explanations for the explananda are constructed and the stronger the resulting best explanation will be (cf. the critical question in Section 4.4.1: how many other stories are there that explain the explananda?). The dialogue game hence allows for two or more players:

**Definition 5.5.2 (Players)**
*The set Players* = $\{p_1, \ldots, p_n\}$ *is a finite set such that* $n \geq 2$ *that includes the players in the dialogue.*

Having two or more players allows us to model, for example, a discussion between prosecution and defence but also discussions in a group of crime analysts or jurors.

The communication language determines which speech acts can be performed in order to influence the theory *AET* (Definition 5.5.3).

**Definition 5.5.3 (Communication Language)**

*The communication language* $L_c$ *is a set of* speech acts. *Each act is of the form* $a(c)$, *where* $a$ *denotes the act and* $c$ *denotes the content of the act. The act types in* $L_c$ *are*:

- *argue* $A$
- *explain* $(E, S)$
- *match* $(S, G_S, G_A)$
- *concede* $\varphi$
- *retract* $\varphi$

*Here, A is an argument, E is a set of ground literals, S is a story,* $G_S$ *is a story scheme and* $G_A$ *is a set of abstraction generalizations.*

The speech acts allow the players to give an argument $A$ (*argue*), explain events in $E$ with a story $S$ (*explain*), predict events given some events $E$, match a story $S$ to a story scheme $G_S$ using some abstraction links expressed by $G_A$ (*match*) and concede to or retract some propositions from their commitments. In the above definition, only the preconditions of the speech acts have been given (i.e. $A$ has to be an argument); in Section 5.5.7 the limiting conditions that determine exactly what can be said with these speech acts will be given.

Given the set of speech acts and players, the moves of a dialogue, which are of the form (*player, speech act*), can be defined. A sequence of moves is called a *dialogue*.

**Definition 5.5.4 (Moves and Dialogues)**

- *The set Moves can be defined as Players* $\times$ $L_c$, *where the elements of a move m are denoted by, respectively:*
  - *pl(m), the player of the move*
  - *sa(m), the speech act performed in the move*
- *The set of finite dialogues D is the set of all finite sequences of moves* $m_1, \ldots, m_n$. *The dialogue* $m_1, \ldots, m_i$ *is denoted with* $d_i$ *and* $d_0$ *stands for the empty dialogue. When d is a dialogue and m a move then d, m denotes the continuation of d with m.*

For example, $d_2 = (p_1$, *argue* (H in car)), $(p_2$, *concede* Haaknat robs the supermarket) is a dialogue consisting of two moves. The combination of this dialogue with a new move can be denoted as $d_2$, $(p_3$, *argue* (H not in car)). Note that the set of dialogues consists of all finite sequences of moves but this does not yet define what a proper dialogue is. That is, the rules of the dialogue still have to be defined in the form of a protocol.

### 5.5.3 The Hybrid Theory in a Dialogue

One of the important features of the current dialogue game is that the players jointly build on their shared knowledge about the case. This shared knowledge has the form of an argumentative explanation theory *AET.* At the start of the dialogue, the set explananda $F$ is assumed to be nonempty (see Definition 5.5.5); the reason for this is that otherwise there would be no explanandum to start the dialogue with.

**Definition 5.5.5 (The Initial Theory)**

*AET*($d$) *denotes the hybrid theory AET after dialogue* ($d$). *The initial theory AET*($d_0$) *is constrained as follows:* $F(d_0) \neq \emptyset$.

Note that because of this constraint, the initial set of evidence and argument must also be nonempty as Definition 5.4.1 requires that every explanandum is supported by a piece of evidence.[14] The set $F$ does not change during the dialogue, so it must be agreed upon before the dialogue starts. It is in theory possible to have an argumentative dialogue about what the explananda are. However, the purpose of the current dialogue is to find explanations for certain observations and to compare these explanations; a dialogue about what should be explained is a different kind of dialogue the details of which we leave for future research.

There are no other constraints on the elements of *AET* at the beginning of the dialogue and after each move by a player, the theory is updated:

**Definition 5.5.6 (The Hybrid Theory in a Dialogue)**

*The* theory update function *is a function U: D* $\rightarrow$ *AET*.

*AET*($d$) *stands for the theory after dialogue d. Individual elements of the theory are denoted similarly,* e.g. *ET*($d$), *Args*($d$).

*AET*($d$, $m$) *can be defined as follows:*

(1) *If* $sa(m) = argue\ A$, *then*
   (a) $G_E(d, m) = G_E(d) \cup Gens(A)$.
   (b) $I_E(d, m) = I_E(d) \cup Evidence(A)$.
   (c) $Args(d, m) = Args(d) \cup \{A\}$.
   (d) *Attack*($d$, $m$) *is updated according to Definition 5.2.6.*
(2) *If* $sa(m) = explain\ (E, S)$, *then*
   (a) $G_C(d, m) = G_C(d) \cup Gens(S)$.
   (b) $H(d, m) = H(d) \cup Events(S)$.
   (c) $Expl(d, m) = (Expl(d) \cup \{S\})$.
(3) *If* $sa(m) = match\ (S, G_S, G_A)$, *then*
   (a) $Schemes(d, m) = Schemes(d) \cup \{G_S\}$
   (b) $G_{CA}(d, m) = G_{CA}(d) \cup G_A$.

[14]However, in the simple examples below the arguments for the explananda will often not be explicitly mentioned.

(4) *If* $sa(m) = retract(\varphi)$, *then*
  (a) *If* $\forall p \in Players$: $\varphi \notin C_p$ *then,*
    (i) $\varphi$ *is removed from its corresponding element in* $CF(d)$ *or* $ET(d)$*; and*
    (ii) *Any argument* $A$ *such that* $\varphi \in A$ *is removed from Args; and*
    (iii) *Any explanation* $S$ *such that* $\varphi \in S$ *is removed from Expl.*

*Elements of the AET which are not mentioned in the above list do not change after the respective move.*

The update function takes as its input a dialogue and gives as output a theory *AET*. Condition 1 says that moving an argument adds its generalizations, evidence and the argument itself to their respective sets in the theory (see Definition 5.2.5 and Definition 5.2.7). Condition 2 says that posing an explanation adds the generalizations, events and the story to their respective sets in the theory (see Definition 5.3.3 and Definition 5.4.1). Condition 3 says that a *match* speech act adds the story scheme to the set *Schemes* and the abstraction generalizations to $G_{CA}$. Finally, condition 4 says that if a player retracts a proposition and through this retraction, none of the players is committed to the proposition, it is deleted from the one of the theories *CF* or *ET* and any argument or explanation that has $\varphi$ as an element is also deleted from the *AET*. Because only those arguments and explanations that have explicitly been moved are added to their respective sets *Args* and *Expl*, for the *retract* move it also has to be specified which arguments or explanations have to be deleted. The concede move does not change the *AET*, as a player can only concede to something which has already been moved in the dialogue before and hence is already in the *AET*.

### 5.5.4 Commitments

The *AET* from Definition 5.5.6 represents the total set of shared knowledge the players have built. However, a player does not have to agree with everything that is part of this shared hybrid theory. For example, if the theory *AET* contains the arguments (*H in car*) and (*H not in car*), it does not make sense that one player is committed to both arguments as in such a case the player would be committed to contradictory knowledge. Therefore, a player also has his own separate commitment store which contains the parts of the theory he is committed to. Note that the commitments of a player are public and that all the players have knowledge of the commitments of the others. A player's commitments are influenced by the moves they play during the dialogue.

**Definition 5.5.7 (Commitment Rules)**
*The* commitment function *is a function* $C: D \times Players \rightarrow \wp(\mathcal{L}_d)$.

$C_p(d)$ *denotes the commitments of player p after dialogue* $(d)$.

*For any player p, the initial commitments are constrained as follows:* $\varphi \in C_p(d_0)$ *iff* $\varphi \in ET(d_0)$ *or* $\varphi \in CF(d_0)$.

*$C_p(d, m)$, where $pl(m) = p$, can be defined as follows:*

(1) $C_p(d, argue\ A) = C_p(d) \cup Prem(A)$.
(2) $C_p(d, explain\ (F, S)) = C_p(d) \cup Gens(S) \cup Events(S)$.
(3) $C_p(d, match\ (S, G_S, G_A)) = C_p(d) \cup G_S \cup G_A$.
(4) $C_p(d, concede\ \varphi) = C_p(d) \cup \{\varphi\}$
(5) $C_p(d, retract\ \varphi) = C_p(d) \setminus \{\varphi\}$

*Player p is* committed to an explanation $S \in Expl(d)$ *iff for all* $\varphi$ *such that* $\varphi \in Events(S_i)$ *or* $\varphi \in Gens(S_i)$, $\varphi \in C_{pi}(d)$.

So the commitment function has as its input a dialogue and a player and as output a set denoting the player's commitments. The players' commitments before the dialogue starts are constrained in that the players are always committed to the propositions which are in the initial *AET* and nothing more. Players automatically commit themselves to all parts of an argument (1), explanation (2) or scheme (3) they move in the dialogue and to any abstraction generalizations they move (3). Conceding to a proposition adds it to a player's commitments (4), while retracting a proposition deletes it from the commitments (5). Finally, a player is committed to an explanation in *Expl* if and only if he is committed to all its events and generalizations.

As an example consider the initial commitments of player $p_1$ to be {police find Haaknat in park}. If $p_1$ now performs the move *explain* (police find Haaknat in park, [Haaknat robs supermarket, police arrives, $g_{c3}$, $g_{c4}$, $g_{c5}$], his commitments after this move will be {Haaknat robs supermarket, police arrives, Haaknat gets into car and takes off, Haaknat thinks police are following, Haaknat parks car, Haaknat hides in park, police find Haaknat in park, police search in park, $g_{c3}$, $g_{c4}$, $g_{c5}$}. In words, $p_1$ has committed himself to all the individual elements of the explanation.

## 5.5.5 Comparing Explanations

Above it was argued that the current winner of the game should be the player who is committed to the explanation that is currently considered the "best". The interesting point here is determining which explanation is currently the best. It was already argued that defining an exact (mathematical) measure of what is the best explanation is not a trivial task. The following definitions are an attempt at providing such a measure. It should be noted that this ordering is only used to determine the *current winner* and thus control turn taking in the dialogue game. The ordering does not provide hard-and-fast rules on what the ultimate best explanation is, as this decision always involves a substantive and context-dependent element. The first way of ordering compares the stories according to set inclusion on the criteria:

**Definition 5.5.8 (Comparing Explanations)**

*Given two explanations $S_i$ and $S_j$, a partial ordering function $\leq$ can be defined as follows:*

- *If* $E^+(S_i) \subset E^+(S_j)$ *and* $E^-(S_i) \supseteq E^-(S_j)$ *then* $S_i < S_j$
- *If* $E^+(S_i) \subseteq E^+(S_j)$ *and* $E^-(S_i) \supset E^-(S_j)$ *then* $S_i < S_j$
- *If* $E^+(S_i) = E^+(S_j)$ *and* $E^-(S_i) = E^-(S_j)$ *then*
  - *If* $P^+(S_i) \subset P^+(S_j)$ *and* $P^-(S_i) \supseteq P^-(S_j)$ *then* $S_i < S_j$
  - *If* $P^+(S_i) \subseteq P^+(S_j)$ *and* $P^-(S_i) \supset P^-(S_j)$ *then* $S_i < S_j$
  - *If* $P^+(S_i) = P^+(S_j)$ *and* $P^-(S_i) = P^-(S_j)$ *then*
    - *If* $E^G(S_i) \supset E^G(S_j)$ *then* $S_i < S_j$
    - *If* $E^G(S_i) = E^G(S_j)$ *then* $S_i = S_j$

*If none of the above conditions hold,* $S_i$ *and* $S_j$ *are incomparable.*

In words, if $S_j$ covers more evidence and is contradicted by less or an equal amount of evidence than $S_i$, then $S_j$ is better than $S_i$. Similarly, if $S_j$ is contradicted by more evidence and covers a less or equal amount of evidence than $S_i$, then $S_j$ is better than $S_i$. If both the evidential support and the evidential contradiction are the same, then the plausibility and the implausibility of the two explanations is checked in the same way as the evidential support and contradiction. If this is also equal, then the evidential gaps are checked. Matching and completeness are not checked here because they are absolute measures (a story either completely matches or does not completely match a scheme); these criteria will be incorporated into Definition 5.5.9 below.

Note that evidential support and contradiction are considered more important than plausibility or evidential gaps. If $S_1$'s evidential support is better than $S_2$'s, $S_1$ will be the better explanation, even if the plausibility of $S_2$ is better than that of $S_1$. This ensures that a good story does not push out a true story. Finally, the definition compares two stories using set inclusion: if, for example, the set denoting the evidential support of $S_2$ is a subset of the set denoting the evidential support of $S_1$, then $S_1$ is better. The idea behind this is that story $S_1$ only wins over story $S_2$ if $S_1$ is supported by at least the same evidence as $S_2$ and then some more. This way of ordering explanations with set inclusion provides a solid criterion for determining which explanation is the best and it ensures that there is no alternative explanation which is actually better.

For current purposes, the main disadvantage of this way of ranking the criteria according to set inclusion is that in this way stories are often incomparable. Another way of defining an ordering on explanations is to consider the absolute degrees of evidential support and contradiction, plausibility and implausibility and evidential gaps. In other words, the absolute evidential support is the number of pieces of evidential data that support the explanation. While this may seem somewhat ad hoc, it is in my opinion a reasonable way of comparing explanations for current purposes (i.e. to determine a current winner). Assuming that in a real case the set of evidence that can be collected and the time that can be spent on discussing the possibilities is finite, it makes sense to require that an explanation explains as much of the evidence as possible. The above Definition 5.5.8 can be modified to allow for this new way of comparing explanations. In this new Definition 5.5.9, evidential support and contradiction are still more important than plausibility. New criteria are those of matching and completeness: story that completely matches a scheme is better than other stories and a complete story is better than an incomplete story. Thus completeness is more important than plausibility, which allows for the more holistic reasoning with stories, where story schemes are judged as being more important than (the plausibility of) individual causal generalizations. An advantage of this way of ordering explanations is that two explanations can always be compared and that thus there is a total ordering on the set of explanations.

**Definition 5.5.9 (Comparing Explanations II)**

*Given two explanations $S_i$ and $S_j$, a total preordering function $\leq_a$ can be defined as follows:*

- *If $|E^+(S_i)| < |E^+(S_j)|$ and $|E^-(S_i)| \geq |E^-(S_j)|$ then $S_i <_a S_j$*
- *If $|E^+(S_i) \leq |E^+(S_j)|$ and $E^-(S_i) > |E^-(S_j)|$ then $S_i <_a S_j$*
- *If $|E^+(S_i)| = |E^+(S_j)|$ and $E^-(S_i) = |E^-(S_j)|$ then*
  - *If $\exists G_{S1} \in$ Schemes such that $S_i$ completely matches $G_{S1}$ and $\neg\exists G_{S2} \in$ Schemes such that $S_j$ completely matches $G_{S2}$ then $S_i <$ S$_j$; otherwise*
  - *If $\exists G_{S1} \in$ Schemes such that $S_i$ is complete w.r.t. $G_{S1}$ and $\neg\exists G_{S2} \in$ Schemes such that $S_j$ is complete w.r.t. $G_{S2}$ then $S_i <$ S$_j$; otherwise*
  - *If $|P^+(S_i)| < |P^+(S_j)|$ and $|P^-(S_i)| \geq |P^-(S_j)|$ then $S_i <_a S_j$*
  - *If $|P^+(S_i)| \leq |P^+(S_j)|$ and $|P^-(S_i)| > |P^-(S_j)|$ then $S_i <_a S_j$*
  - *If $|P^+(S_i)| = |P^+(S_j)|$ and $|P^-(S_i)| = |P^-(S_j)|$ then*
    - *If $|E^G(S_i)| > |E^G(S_j)|$ then $S_i <_a S_j$*
    - *If $|E^G(S_i)| = |E^G(S_j)|$ then $S_i =_a S_j$*

*Where $|S|$ stands for the number of elements of set S.*

The combination of the above two definitions can be used to provide measures with which the critical question "how decisively does the current story surpass the alternative stories?" (see Section 4.4.1) can be answered. Even though often the decision as to what constitutes the best explanation involves a substantive and context-dependent element, there are cases in which it is clear that one of the two cases is clearly the best. If, for example, story $S_1$ fares better (w.r.t. set inclusion) than $S_2$ on evidential support, evidential contradiction, completeness and plausibility, we can reasonably safely say that $S_1$ is the best explanation. Another example is when out of the 20 pieces of evidence, story $S_1$ is supported by 18 of these and story $S_2$ is contradicted by 18. In other words, in less complex cases where the difference in quality between the available explanations is significant, the above definitions can quite safely be used to determine which explanation is the best. In complex cases where no explanations really stands out, the above definitions and criteria serve only to define heuristics that point to possible points where the opponents' explanation can be attacked or one's own explanation can be improved.

## *5.5.6 Current Winner and Turntaking*

One way in which the above orderings can be used is to determine the *current winner* in a dialogue (Definition 5.5.10). The current winner is a function that determines the player or players who are committed to the best explanation according to the above orderings. Note that the current winner is not the *final winner* of the dialogue but rather the player which is in the best position at a certain point in a (non-terminated) dialogue.

**Definition 5.5.10 (Current Winner)**

*The* winning function *is a function Winner:* $D \rightarrow \wp(Players)$.

*Winner*$(d)$, *the current winners of dialogue d, can be defined as follows:*

*Player* $p_i \in$ *Winner(d) iff*

(1) *either*

  (a) *There is an explanation* $S_i \in Expl(d)$ *such that for each explanation* $S_j \in Expl(d)$ (*where* $S_i \neq S_j$), *it holds that* $S_j < S_i$; *and*
  (b) $p_i$ *is committed to* $S_i$;

(2) *or*

  (a) *There is* ***no*** *explanation* $S_i \in Expl(d)$ *such that for each explanation* $S_j \in Expl(d)$ (*where* $S_i \neq S_j$), *it holds that* $S_j < S_i$; *and*
  (b) *there is an explanation* $S_k \in Expl(d)$ *such that for each explanation* $S_l \in Expl(d)$ (*where* $S_k \neq S_l$), *it holds that* $S_l <_a S_k$; *and*
  (c) $p_i$ *is committed to* $S_k$.

Given a dialogue, a player is a current winner if he is committed to an explanation in *AET* that is better than the other explanations in *AET* according to Definition 5.5.8. If there is no such explanation, the players that are committed to the best explanation according to Definition 5.5.9 are the current winners. If there is no such explanation, there is no current winner. Note that there can be more than one current winner because more than one player can be committed to the best explanation simultaneously. As an example, say that the first player moves the explanation [Haaknat robs supermarket, police arrives, $g_{c3}$, $g_{c4}$, $g_{c5}$] from Section 5.3.2. This player is then the winner, because he is the only one that is committed to an explanation.

The definition of a current winner is mainly used to control *turntaking*, that is, to determine which who's turn it is to make a move in the dialogue (Definition 5.5.11 on p. 151).

**Definition 5.5.11 (Turntaking)**

*The* turntaking function *is a function Turn:* $D \rightarrow$ *Players, where Players* $= \{p_1, \ldots, p_n\}$.

*Turn*$(d)$, *the player to move in dialogue d, can be defined as follows:*

– $Turn(d_0) = \{p_1\}$
– $Turn(d, m) = p_i$ *if* $Turn(d) = p_i$ *and* $Winner(d) \neq p_i$

– *Turn*($d$, $m$) = $p_{i+1}$ *if Turn*($d$) = $p_i$ *and Winner*($d$) = $p_i$ *and* $i < n$
– *Turn*($d$, $m$) = $\{p_1\}$ *if Turn*($d$) = $p_i$ *and Winner*($d$) = $p_i$ *and* $i = n$

The turntaking function has as its input a dialogue and as its output the player whose turn it is to move. When the dialogue starts, there is no winner so the first player starts the dialogue. After this, if a player becomes a current winner by making a move, the turn goes to the next player. If this player is already a winner, he passes on the turn to the next player. So each player keeps his turn until he is able to become a current winner. Take the above example: after the first player poses the explanation, he becomes the current winner so the turn goes to the second player. In order to become a winner, this player should also commit to an explanation (either by moving one himself or by committing to the first player's explanation). Note that if the second player merely attacks the first player's explanation, he does not become a winner because he is not committed to an explanation. So moving an argument, while perhaps permitted, does not change the second player's status as a winner and hence does not change the turn to the next player.

## *5.5.7 The Protocol*

The protocol *P* specifies the allowed moves at each stage of a dialogue. Essentially, the protocol defines the rules of the game and is meant to ensure that the players' moves are relevant to the inquiry. First, the general form of a protocol function should be defined (adapted from Prakken, 2005b):

**Definition 5.5.12 (Protocol Function)**

*A* protocol *on Moves is a set* $P \subseteq D$ *such that whenever dialogue* $d \in P$, *then all initial sequences of* $d \in P$.

*A* protocol function *is a partial function* $Pr: D \rightarrow \wp(Moves)$ *such that*

– $Pr(d)$ *is undefined if* $d \notin P$
– $Pr(d) = \{m \mid d, m \in P\}$ *otherwise.*

*The elements of the domain of Pr are the* proper finite dialogues.

So the protocol function $Pr(d)$ determines what moves are allowed after dialogue $d$ and if $d$ is not a proper dialogue, then no protocol function is defined for the dialogue. We can now also assume that a protocol must satisfy some conditions, so that only some particular types of moves may be allowed, given the current state of the dialogue.

**Definition 5.5.13 (Protocol Conditions)**

*For all moves m and proper finite dialogues* $d_i$ *it holds that* $m \in Pr(d_i)$ *if and only if all of the following conditions are satisfied:*

(1) *If there is no move* $m_i \in d_i$ *such that* $pl(m_i) = pl(m)$ *then m is an explain* $(F, S)$ *move such that there is no substory of S in* $Expl(d_i)$.

(2) $pl(m) = Turn(d_i)$
(3) *m was not already moved in d by the same player*
(4) $C_{pl(m)}(d_i, m) \nvdash \bot$
(5) *If m is an argue A move (where* $\varphi$ *is A's conclusion), then* $A \notin Args(d_i)$ *and*

(a) *A attacks some argument B in Args and* $Conc(B) \cap F = \emptyset$*; or*
(b) *A contradicts some* $\varphi \in CF(d_i)$ *and* $\varphi \notin F$*; or*
(c) $\exists S, S \in Expl(d_i)$*,* $pl(m)$ *is committed to S and* $E^+(S)(d_i, m) > E^+(S)(d_i)$*; or*
(d) $\exists S, S \in Expl(d_i)$*,* $pl(m)$ *is committed to S and* $P^+(S)(d_i, m) > P^+(S)(d_i)$*; or*
(e) $\exists S, S \in Expl(d_i)$*,* $pl(m)$ *is committed to S and* $E^G(S)(d_i, m) < E^G(S)(d_i)$*.*

(6) *If m is an explain (E, S) move, then*

(a) $F \subseteq E$*; and*
(b) $S \notin Expl(S)$*; and*
(c) *it is not the case that* $pl(d_{i-1}) = pl(m)$ *and* $sa(d_{i-1}) = explain\ (F, S')$

(7) *If m is a match* $(S, G_S, G_A)$ *move, then*

(a) $pl(m)$ *is committed to S; and*
(b) $\neg\exists G_S' \in Schemes(d_i)$ *such that* $G_S'$ *is a proper scheme and S completes* $G_S'$ *and*
(c) $G_S$ *is a proper scheme and S completes* $G_S$ *after dialogue (d, m).*

(8) *If m is a concede* $\varphi$ *move, then*

(a) $\varphi$ *is an element of* $CF(d) \cup ET(d)$*; and*
(b) $\varphi \notin C_{pl(m)}(d)$

(9) *If m is a retract* $\varphi$ *move, then*

(a) $\varphi \in C_{pl(m)}(d)$*; and*
(b) $\varphi \notin F$ *or* $\varphi$ *is not in some argument B such that* $Conc(B) \cap F \neq \emptyset$*.*

(10) $\neg\exists S_i$: $S_i \in Expl(d)$ *such that*

(a) *For all* $p_i \in Players$*:* $p_i$ *is committed to* $S_i$*; and*
(b) *for each other explanation* $S_j$*,* $S_j \in Expl(d)$ *and* $S_i \neq S_j$*, it holds that* $S_i >_a S_j$*.*

The protocol basically has as its input a dialogue and outputs a list of admissible moves. The first condition governs the first move of a player in the dialogue, which must be an explain move that is not based on an explanation that has already been moved by another player. This ensures that each player has his own explanation. The next two conditions say that only the player-to-move can make allowed moves and that a player may not repeat his moves. Condition 4 regulates that a player's commitments may not be inconsistent, that is, a player may not contradict himself in the strict logical sense.[15]

Condition 5 states when an argument may be moved, namely if was not already moved before and if it is somehow relevant to the current *AET*; this last condition is enforced by the requirement that (at least) one of the conditions 5a–5e holds. Condition 5a states that an argument may be moved if it attacks another argument which is not an argument for an explanandum. This allows players to reason about the individual evidence, for example, the veracity of a witness. Condition 5b states that an argument may be moved if it contradicts an element in *CF* (except for an explanandum). This condition allows the players to attack the other players' explanations, schemes and correspondence links and this ensures that both types of contradicting, evidential contradiction and implausibility, can be done. Conditions 5c–5e allow a player to give arguments that improve the evidential support or plausibility or decrease the evidential gaps of one of his explanations. The conditions 5a–5e ensure that an argument either attacks the opponent or supports the player's position.

Condition 6 governs the admissibility of *explain* moves. The first condition (6a) ensures relevance of explanations by requiring that the explanation actually explains at least the explananda according to Definition 5.3.4 (6a). Notice that with this condition on explain moves, it is also possible to *predict* new possible observations: as long as the explananda are part of the story, the story can contain any number of additional observations. A disadvantage of this condition is that every time a player wants to change an existing explanation (by, for example, further explaining initial events, predicting new observables or refining a causal relation) the player should move a totally new explanation consisting of the old explanation with the required changes. The old explanation is still part of the set *Expl* and thus this set can quickly get very large as only minor changes to explanations require the addition of new explanation. However, because the current theory is not directly meant to be used in an automated reasoning system, this is not a big disadvantage. Condition 6b says that an explanation may only be moved if the explanation has not been moved before and condition 6c ensures that the previous move was not an *explain* move by the same player (6c). This last condition is added so that players do not simply keep giving explanations without trying to make these explanations the best by improving them or worsening the opponents' explanations. Condition 7 says that a player may only match a story to a scheme if he is committed to that story (7a), the story does not already complete some proper scheme (7b) and the story completes the proper scheme $G_S$ in the move (7c). In other words, if a story does not already complete some proper scheme this can be done with the *match* move.

Condition (8) ensures that a player concedes a proposition only if it is in the current theory and the player is not already committed to it. Condition (9a) says

---

[15] Note that a player may be committed to two arguments that attack each other or to two alternative explanations for the explananda.

that a player can only retract a proposition to which he is committed and condition 9b ensures that a player does not retract an explanandum or a part of an argument supporting an explanandum. Finally, condition (10) implies that a dialogue terminates if all players are fully committed to the best explanation, that is, the dialogue terminates when all the players are winners according to Definition 5.5.10. This essentially means that when the dialogue terminates, the players agree that the inquiry has gone on for long enough and that all possible and sensible options have been considered.

The above protocol ensures that all moves are relevant to the current *AET* and encourages the players to, for example, improve their explanations (conditions 5c and d). Because there are concede and retract moves and it is assumed that at some point the dialogue should terminate, the dialogue does not ensure that tunnel-vision is completely avoided. However, when we assume that the players have the aim to (at least at the start of the dialogue) become the winner, we can say that the above protocol, winning and turntaking conditions together enforces a standard of rationality for the hybrid theory.

## 5.6 An Example of the Dialogue Game

In this section an example of the dialogue is given to illustrate the dialogue game. The example is given formally but also rendered in the "semi-formal" visualisation that is used throughout this book. For the example we return to the Haaknat case. In the example, there are two players and the set of explananda $F=$ {police find Haaknat in park}; the argument for this explanandum will be implicitly assumed. According to Definition 5.5.7, both players are committed to this explanandum. Player $p_1$ starts the dialogue by providing an explanation for this explanandum:

$p_1$: *explain* ($F$, $S_{\text{p1a}} =$ [Haaknat robs supermarket, $g_{\text{c1}}$: Haaknat robs supermarket $\Rightarrow_C$ Haaknat flees, $g_{\text{c2}}$: Haaknat flees $\Rightarrow_C$ Haaknat hides in a park, $g_{\text{c3}}$: Haaknat hides in park $\Rightarrow_C$ Haaknat in park])

The events and generalization from the above story are added to $p_1$'s commitments. Now $p_1$ is winning, because he is committed to the only explanation for $F$. $p_2$ at this point only has two options: he either concedes to the above explanation or he has to provide an explanation for $F$ which is somehow better than $p_1$'s explanation. Assume that $p_2$ provides an explanation for $F$.

$p_2$: *explain* ({police find Haaknat in park}, $S_{\text{p2a}} =$ [argument between Haaknat and Benny, $g_{\text{c4}}$: argument between Haaknat and Benny $\Rightarrow_C$ Haaknat flees, $g_{\text{c2}}$, $g_{\text{c3}}$])

After providing this explanation, it is still $p_2$'s turn, as the explanation he has provided is not better than $p_1$'s explanation $S_{\text{p1a}}$ according to one of the definitions from Section 5.5.5. $p_2$ has a number of options: he can, for example, try to support $S_{\text{p2a}}$, increasing the evidential support or plausibility, or he can try to contradict $S_{\text{p1a}}$, thus increasing the evidential contradiction or implausibility of his opponent's explanation. In the example, $p_2$ supports $S_{\text{p2a}}$ by providing an argument based on evidence.

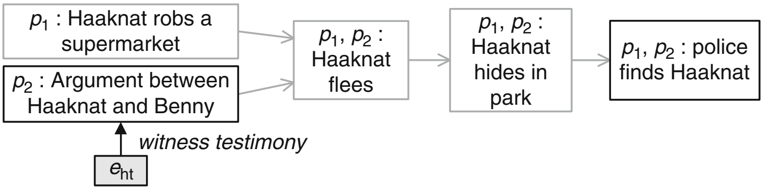



**Fig. 5.8** The *AET* after dialogue $d_3 = (p_1{:}explain\ (F, S_{p1a})), (p_2{:}explain\ (F, S_{p2a})), (p_1{:}argue\ A_{eht})$

$p_2$: *argue* $A_{eht}$:

(1) $e_{ht}$: Haaknat's testimony "I had an argument with Benny". **($I_E$)**
(2) argument between Haaknat and Benny **(*1, witness testimony*)**

Now $p_2$ is the current winner: there is one piece of evidence in *AET* and it supports $S_{p2a}$ so from Definition 5.5.8 it follows that $E^+(S_{p1a}) \subset E^+(S_{p2a})$. The current theory is pictured in Fig. 5.8. For each event, it is indicated which players are committed to that event.

At his point $p_1$ has multiple options. He can, for example, provide an argument for $\neg valid(g_{c4})$. Contradicting $g_{c2}$ and $g_{c3}$, even though this would increase the evidential contradiction or implausibility of $S_{p2a}$, is not an option because then $p_1$ would contradict himself. Just contradicting $g_{c4}$ with an argument based on the stock of knowledge $G_E$ is not enough: recall that evidential support is more important than (causal) plausibility. Another option is to decrease the evidential support of $p_2$'s explanation by defeating the argument $A_{e1}$. For now, $p_1$ chooses to increase the evidential support of his own explanation. Recall from the earlier Haaknat examples that a witness saw Haaknat get into a car near the supermarket just after it was robbed. $p_1$ can try to incorporate this into a direct argument for the fact that Haaknat robbed the supermarket. For example, the ultimate probandum Haaknat robs supermarket can be supported by a complex evidential argument (see Fig. 3.15, p. 55). However, as was argued before, this leads to contrived generalizations. It is in this case easier to extend the story $S_{p1a}$ so that the event Haaknat gets in car is part of it; in other words, the causal link represented by Haaknat robs supermarket $\Rightarrow_C$ Haaknat flees should be refined (see Fig. 3.18, p. 63) into the chain Haaknat robs supermarket $\Rightarrow_C$ Haaknat gets in car, Haaknat gets in car $\Rightarrow_C$ Haaknat flees. There is no special speech act for refining generalizations,[16] but $p_1$ can provide a new explanation with the refined causal chain in it:

$p_1$: *explain* ($F$, $S_{p1b}$ = [Haaknat robs supermarket, $g_{c5}$: Haaknat robs supermarket $\Rightarrow_C$ Haaknat gets in car, $g_{c6}$: Haaknat gets in car $\Rightarrow_C$ Haaknat flees, $g_{c2}$, $g_{c3}$])

This new explanation can be supported by the argument (*H in car*) from p. 114. Figure 5.9 visualizes the new explanation (here $S_{p1a}$ and $A_{e1}$ have not been rendered).

[16] See (Bex and Prakken, 2004) for a way to integrate such a speech act in a dialogue game similar to the current game.

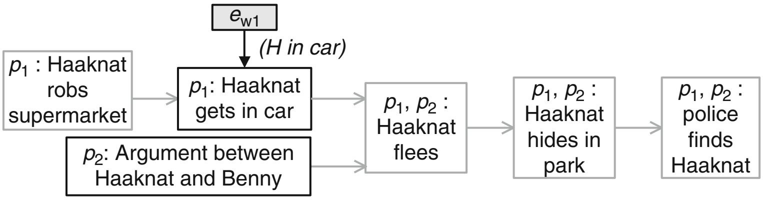



**Fig. 5.9** The *AET* after dialogue $d_5 = d_3$ ($p_2$:*explain* ($F$, $S_{\text{p2b}}$)), ($p_2$:*argue* (H in car))

After the above *explain* move, it is still the case that $S_{\text{p2a}} > S_{\text{p1b}}$ because there are more evidential gaps in $S_{\text{p1b}}$. Thus Definition 5.5.8 ensures that the players do not simply extend their story without providing any new evidence or plausibility. $p_2$ is still the winner and $p_1$ has to make more moves in order to make his new explanation $S_{\text{p1b}}$ better than $S_{\text{p2a}}$. One option for $p_1$ is to try and match his new explanation to the intentional action scheme (Fig. 5.10): if $S_{\text{p1b}}$ matches this scheme or at least completes it, it is better than $p_2$'s explanation $S_{\text{p2a}}$ because matching and completeness are more important then evidential gaps in Definition 5.5.9.

If the central action is Haaknat robs supermarket, then the scheme is not yet fully completed, because a motive and a goal for Haaknat's robbery is still missing. In this way, schemes can also point to possible new events that should be abduced: $p_1$ should further extend explanation $S_{\text{p1b}}$ by explaining exactly why Haaknat robbed the supermarket. Recall that, because of condition 6a of Definition 5.5.13, moves that only "extend" an explanation by are impossible in the formal dialogue game because the explananda always have to be explained. As was shown above in the case of refining, the option available to the players in such a case is to simply provide a new explanation that incorporates the extension or refinement:

$p_1$: *explain* ($F$, $S_{\text{p1c}}$ = [Haaknat is drug addict, $g_{c7}$: $x$ is drug addict $\Rightarrow_C$ $x$ wants to rob supermarket, $g_{c8}$: Haaknat wants to rob supermarket $\Rightarrow_C$ Haaknat robs a supermarket, $g_{c5}$, $g_{c6}$, $g_{c2}$, $g_{c3}$])

This explain speech act is allowed since $p_1$'s last move was an *argue* move. Notice that in the above explanation, one of the generalizations ($g_{c7}$) is given in a more abstract form "drug addicts (generally/usually/sometimes) want to rob supermarkets"; in the actual derivation, $x$ will be instantiated with Haaknat but the generalization is given in its general form so that it is clear at which level of specificity $p_1$ intends the generalization to be. $p_1$ could have also given the more specific

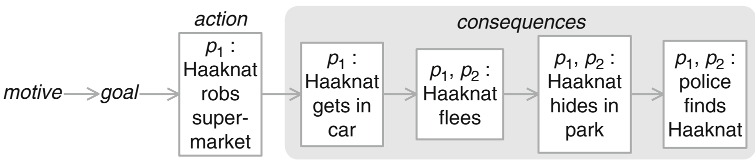



**Fig. 5.10** Trying to match $S_{\text{p1b}}$ to the intentional action scheme

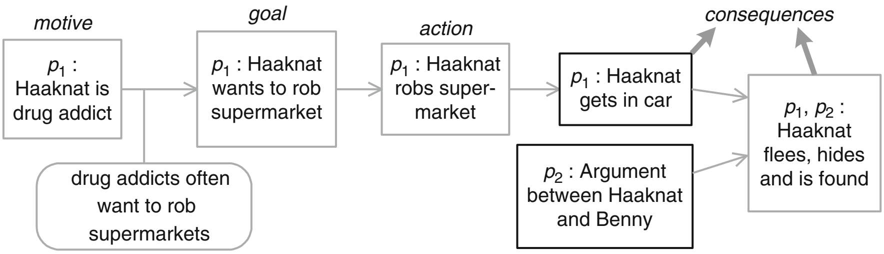



**Fig. 5.11** The *AET* after dialogue $d_7 = d_5$, ($p_1$:*explain* (*F*, $S_{\text{p1c}}$)), ($p_1$:*match* ($S_{\text{p1c}}$, $G_{Sia}$, $G_{A1}$)

form Haaknat is drug addict $\Rightarrow_C$ Haaknat wants to rob supermarket, which can be paraphrased as "Haaknat's drug addiction will cause him to rob a supermarket". Now the new story $S_{\text{pic}}$ can be matched to the intentional action scheme:

$p_1$: *match* ($S_{\text{p1c}}$, $G_{Sia}$ = {motive $\Rightarrow_C$ goal, goal $\Rightarrow_C$ action, action $\Rightarrow_C$ consequence}, $G_{A1}$ = {$g_{a1}$: Haaknat is drug addict $\Rightarrow_A$ motive, $g_{a2}$: Haaknat wants to rob supermarket $\Rightarrow_A$ goal, $g_{a3}$: Haaknat robs supermarket $\Rightarrow_A$ action, $g_{a4}$: Haaknat gets in car $\Rightarrow_A$ consequence, $g_{a5}$: Haaknat flees $\Rightarrow_A$ consequence, $g_{a6}$: Haaknat hides in park $\Rightarrow_A$ consequence, $g_{a7}$: police find Haaknat $\Rightarrow_A$ consequence}

A technicality here is that, according to Definition 5.5.6, the explanations $S_{\text{p1a}}$, $S_{\text{p1b}}$ and $S_{\text{p1c}}$ are three separate explanations in *AET*. However, often when one explanation is a substory of another explanation, only the best explanation will be shown in a figure.

In Fig. 5.11, $S_{\text{p1c}}$ matches the intentional action scheme and is equally well supported by evidence as $S_{\text{p2a}}$. The turn goes back to $p_2$, who must try to improve his own explanation or somehow make $S_{\text{p1c}}$ less plausible. For example, if $p_2$ manages to contradict the generalization $g_{c7}$ with hard evidence (for example research that shows that drug addicts commit just as many crimes as the average person), the evidential contradiction of $S_{\text{p1c}}$ goes up and the turn switches back to $p_1$ (because $S_{\text{p2a}}$ is again better than $S_{\text{p1c}}$). However, it is probably hard to find such evidence. $p_2$ can also attack $g_{c7}$ with an argument from the stock of knowledge $G_E$:

$p_2$: *argue* $A_{e2}$:

(1) $i_{da}$: That drug addicts often want to rob supermarkets is based on prejudice. (***$G_E$***)
(2) $g_{e1}$: That drug addicts often want to rob supermarkets is based on prejudice $\Rightarrow_E$ $\neg$*valid*($g_{c7}$) (***$G_E$***)
(3) $\neg$*valid*($g_{c7}$) (***1, 2, DMP***)

However, even after moving this argument $p_2$ is still is not the winner because $S_{\text{p1c}}$ completes the intentional action scheme and this is more important than plausibility. So $p_2$ also has to make his own explanation complete by completing and fitting a proper scheme. First $p_2$ extends his explanation with a reason for why he and

Bennie were in an argument and with the fact that Haaknat was scared because of the argument so he fled:

$p_2$: *explain* ($F$, $S_{\text{p2b}}$ = [Haaknat lends Bennie money, $g_{\text{c8}}$: Haaknat lends Bennie money $\Rightarrow_{\text{C}}$ Bennie does not want to give money back, $g_{\text{c9}}$: Bennie does not want to give money back $\Rightarrow_{\text{C}}$ argument between Haaknat and Bennie, $g_{\text{c10}}$: argument between Haaknat and Bennie $\Rightarrow_{\text{C}}$ Haaknat feels threatened, $g_{\text{c11}}$: Haaknat feels threatened $\Rightarrow_{\text{C}}$ Haaknat flees, $g_{\text{c2}}$, $g_{\text{c3}}$])

Not wanting to try and match his story to the intentional action scheme, $p_2$ comes up with his own scheme on arguments about money and matches his story to that:

$p_2$: *match* ($S_{\text{p2b}}$, $G_{Sl}$ = {argument about money between *x* and *y* $\Rightarrow_{\text{C}}$ *x* feels threatened, *x* feels threatened $\Rightarrow_{\text{C}}$ *x* flees, *x* flees $\Rightarrow_{\text{C}}$ *x* hides, *x* hides $\Rightarrow_{\text{C}}$ *x* is found}, $G_{A2}$ = {$g_{\text{a8}}$: Haaknat lends Bennie money $\wedge$ Bennie does not want to give money back $\wedge$ argument between Haaknat and Bennie $\Rightarrow_{\text{A}}$ argument about money between Haaknat and Bennie, $g_{\text{a12}}$: Haaknat hides in park $\Rightarrow_{\text{A}}$ Haaknat hides, $g_{\text{a13}}$: Police find Haaknat $\Rightarrow_{\text{A}}$ Haaknat is found})

The scheme that $p_2$ gives here tells us that people lend each other money and that this can lead to arguments. Arguments lead to threats and these lead to persons fleeing and hiding from the persons who threaten them. Notice that the components *x* feels threatened and *x* flees of the scheme are not connected to the story through an abstraction generalization because *x* can simply be instantiated with Haaknat. Now the total *AET* can be represented as in Fig. 5.12 (again, only the best explanation for each player has been represented). In the situation represented in this figure, $p_2$ is the winner. $p_1$ is committed to $S_{\text{p1a}}$, $S_{\text{p1b}}$, and $S_{\text{p1c}}$ which are all worse than $S_{\text{p2b}}$: $S_{\text{p1a}}$ has lower evidential support, $S_{\text{p1b}}$ is not complete and $S_{\text{p1c}}$ has lower plausibility.

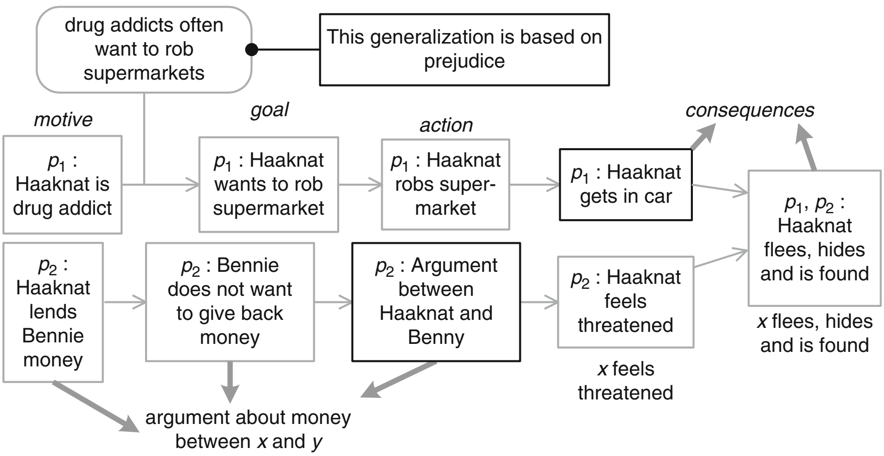



**Fig. 5.12** The *AET* after dialogue $d_8 = d_6$, ($p_2$:*scheme* $G_{\text{S1}}$), ($p_2$:*explain* $S_{\text{p2b}}$)

Recall that in the Haaknat case, the fact that the supermarket was robbed was accepted without a doubt. Now, it can be argued that supermarket is robbed should therefore be an explanandum. However, the subject of the current dialogue is not why and by whom the supermarket was robbed but rather why Haaknat was in the moat. The event supermarket is robbed can, however, be used by $p_1$ to increase his evidential support. First, the argument for this event has to be given:

$p_1$: *argue* $A_{e1}$:

(1) $e_{pr}$: police report that the supermarket was robbed. ($\boldsymbol{I_E}$)
(2) supermarket is robbed (***1, document***)

Now $p_1$ can use the *explain* speech act to predict the event supermarket is robbed through an abstraction link:

$p_1$: *explain* ($F \cup$ {supermarket is robbed}, $S_{p1d} = [S_{p1c}$, Haaknat robs supermarket $\Rightarrow_A$ supermarket is robbed])

Player $p_1$ effectively improves his explanation by performing predictive reasoning towards supporting evidence: a new event is predicted which is supported by evidence. Now $p_1$ is winning because his new explanation $S_{p1d}$ has a higher evidential support than any of $p_2$'s explanations. That the plausibility is lower no longer matters: an implausible story supported by evidence is better than a plausible story that is not supported by evidence. Note that here $p_1$ could just as well have chosen his "older" explanation $S_{p1b}$. This explanation was not complete but less implausible than $S_{p1c}$. In that case, $p_1$ would still be the winner even though his explanation is incomplete, as the fact that someone had a motive and a goal is less important if there is more evidence supporting the explanation that is about the actions.

The explain speech act can also be used to predict *causal story-consequences*. For example, $p_2$ can predict that the argument between Haaknat and Bennie, which according to Haaknat was in a public place, must have caused some witnesses to see the two men argue (p. 156).

$p_2$: *explain* ($F \cup$ {someone must have seen Haaknat and Bennie argue}, $S_{p2c} = [S_{p2b}$, argument between Haaknat and Bennie $\Rightarrow_C$ someone must have seen Haaknat and Bennie argue])

Player $p_2$ should now support this new predicted event, thus increasing the evidential support of his explanation. Summarizing, predictive reasoning towards supporting evidence can be done both with abstraction and causal generalizations.

One use of predictive reasoning which is not modelled in the current dialogue game is predicting story-consequences which are subsequently contradicted by evidence. For example, assume the situation is as it was after dialogue $d_8$ (Fig. 5.12). Here it would be a sensible move for $p_1$ to first extend $p_2$'s explanation (that Haaknat and Bennie has an argument) by predicting someone must have seen Haaknat and Bennie argue and subsequently giving an argument against someone must have seen Haaknat and Bennie argue. In this way, $p_1$ effectively says that "given your story (that Haaknat and Bennie has an argument), someone must have seen them argue. However, all the witnesses who were at the location where Haaknat and Bennie supposedly argued did not see Haaknat or Bennie". The problem for the formal dialogue game here is that $p_1$ changes an explanation to which $p_2$ is committed. Because an explain move by $p_1$ cannot change $p_2$'s commitments, $p_2$ is not committed to the new, contradicted explanation and that $p_2$'s position in the dialogue is therefore not changed.

## 5.7 Evaluation

In this chapter, it has been shown that the hybrid theory of reasoning with arguments and stories can be formalized as a combination of formal argumentation and abductive-causal, model-based reasoning. Section 5.1 introduced a general defeasible logic for the hybrid theory, in which generalizations and their validity play an important role. In Section 5.2, specific evidential prima facie reasons, which capture the context of argumentative reasoning with evidence, have been added to this logic. Section 5.2 continued with definitions of arguments, sequences of inferences based on evidential data or general knowledge, and attacks between arguments. These attacks can have two forms: rebutting attacks negate an explicit line of argument and undercutting attacks deny an inference that is used to infer one line of argument from the other. Based on (Dung, 1995), various argument-based semantics for determining the status of arguments that attack and defeat each other have also been defined.

Section 5.3 defined stories as sequences of events connected by causal relations which are expressed by causal generalizations. Together with abstraction generalizations, a story can then *explain* a set of events if the events are part of the story, that is, if the events explicitly causally follow from the story. Stories that provide different explanations for some events are considered to be alternative explanations and have to be compared according to the criteria further defined in Section 5.4. Finally, in Section 5.3.4 it has been shown how events in stories or sequences of events in stories can correspond to a (part of) a story scheme through abstraction links.

In Section 5.4, the combination of the evidential argument-based part and the causal story-based part, the hybrid theory, has been defined. This combination is best expressed by the various criteria that govern the quality of a story: the *evidential support* and *evidential contradiction* are the pieces of evidence that support or contradict a story through arguments based on evidence. The *evidential gaps* are the events in a story for or against which there is no evidence. The *plausibility* and *implausibility* of a story concerns its evidential gaps and is hence established independently from the evidence. Arguments from general knowledge can be used to discuss this plausibility or implausibility of events and causal generalizations for or against which there is no direct evidence, thus allowing for the pursuit of a cognitive consensus about the general knowledge underlying a story. Finally, a story is *complete* with respect to a scheme if for every element of the scheme it has a corresponding state or event and if all elements of the story correspond to some element in the scheme. This completeness is only relevant if it is with respect to a plausible story scheme, that is, a scheme which is not contradicted by general knowledge.

Section 5.5 proposed a detailed dialogue game in which the players together build a hybrid theory consisting of stories and arguments. The game has an adversarial side in that the players the aim to become the current winner and hence try to expand and improve their own explanation whilst at the same time contradicting the other players' explanations. The protocol rules for the dialogue game ensure that the moves which are made are relevant and that the players' moves are geared towards either improving their own explanation or disproving the opponents' explanations.

The logical theory has been developed to underpin a sense-making tool. That is, the theory has conceptual, cognitive as well as computational aims. This is reflected by the modelling choices made in this chapter, which are not mainly aimed towards computability (as is common in AI) but also towards correctly modelling the concepts that play a part in the process of proof and towards providing a logical that is understandable for people with less formal mathematical training. The conceptual aims of the theory have in my opinion clearly been met. The formal theory has disambiguated between the important concepts of argument and story, clarified the various criteria for judging the quality of stories (in particular the criteria of plausibility and completeness) and, through the dialogue game, provided a proper formal basis for a rational process of proof. The theory formally models ideas from legal theory and legal psychology, such as the combination of arguments and stories and the process of an inquiry, and thus also satisfies any cognitive aims. The weakest part of the formal theory is perhaps its computational part. Some of the aspects of the dialogue game (such as the fact that even for a small change in a story a totally new explanation must be moved) are computationally inefficient. Furthermore, no formal proofs of the computational properties of the logical theory have been given.

One of the main additions of the logical theory to existing formal-logical research is that it provides a combined argumentative and causal model-based approach. The use of causal models allows for abductive reasoning with complex causal models. The use of arguments to discuss the quality of a causal model and the use of hierarchies of causal models (i.e. stories and story schemes) in determining the best explanation are improvements over traditional theories of abductive model-based reasoning. In other model-based approaches, the (causal) model is given and cannot be discussed; using formal argumentation, reasoners can reach a cognitive consensus about the causal model of a case that is compatible with the evidence in a natural and rational way.

The formally defined criteria that determine the quality of a story can be seen as answers to the critical questions from the previous chapter (see Section 4.4). Thus, when these questions are asked in a dialogue they can be answered in detail, which improves the discussion about evidence and proof. The adversarial dialogue game further enforces a rational discussion by allowing player to improve their own explanation and disprove other explanations and by encouraging that, if more than one player participates, multiple explanations are put forward. As far as I am aware, the dialogue game is the first formal game for an inquiry dialogue; the reason for this might be that inquiry is a complex dialogue type which combines both explanation and argumentation.

The formal theory presented in this chapter presents a solid basis for the conceptual (informal) hybrid theory of arguments and stories. However, the theory may be further enriched to allow for other modes of reasoning in the process of proof. One possible to the formal theory is reasoning with *time*. Reasoning with timed events and time intervals can play an important role in the process of proof (for example, "given that the suspect was seen leaving his house at 8:00, is it possible that the suspect was the person who murdered the victim at the other side of town at 8:30?"). Examples of well-known formal models of reasoning with timed events in AI are Kowalski and Sergot's (1986) Event Calculus and the work by Allen and Ferguson (1994); the ideas presented in this and other work on temporal reasoning can have important implications for the current model of stories.

While the formal theory provides the above results and improvements over the current state-of-the-art research, the dialogue game has one conceptual weakness, namely that it does not allow for the contradiction of predicted story consequences (i.e. "from your story $p$ should also causally follow, but I have evidence for $\neg p$"). While such reasoning is in principle possible in the formal hybrid theory, the current definition of the commitments of players does not allow one player to change another player's commitments and hence one player cannot force the other player to accept the predicted event that is contradicted by evidence. Given the importance of such predictive reasoning towards contradictory evidence in the process of proof, it should ideally be incorporated in the formal dialogue game.

Another interesting addition to the dialogue game would be to devise strategies that the players could follow in the dialogue game. For example, in a game with multiple players one player could take on the role of the creative thinker who constantly proposes and extends new explanations whilst another player could take on the role of the sceptic who mainly tries to find fault with other players' ideas and explanations. Various combinations of these strategies could lead to different kinds of dialogues and thus different results. For example, a dialogue with just creative players would be good to avoid tunnel-vision, whereas the strategies of a sceptic player are more interesting for the counsel for the defence.

# Chapter 6
# Case Study: Murder in Anjum

On Christmas Eve 1997, the police found the body of a man with a fractured skull in the front yard of a boarding house in Anjum, a village in the north of the Netherlands. The next day the decomposed remains of another man were found on the grounds of the boarding house. The local media quickly gained interest in the case and the case also attracted some nationwide attention. After the investigation the proprietor of the boarding house, Marjan van der E., was convicted of two murders and sentenced to 6 years imprisonment and detention under hospital order by the Court of Appeals in Leeuwarden. Until this day, Marjan denies any involvement in the murders. In their book, "Moord in Anjum" (Murder in Anjum), Crombag and Israëls (2008) give detailed overviews of the evidence and of the persons involved in the case. Crombag and Israëls, together with other members of the project "Gerede Twijfel" (*Reasonable Doubt*) of Maastricht University, examined most of the original files of the case, visited Anjum and talked to Marjan van der E. In addition to the judiciary's view on the case, they provide a number of alternative scenarios of what possibly happened in Anjum and compare these scenarios using a slightly simplified version of the Anchored Narratives Theory (Wagenaar et al., 1993).

In this chapter, one part of the Anjum case, namely the murder of Leo de Jager,[1] will be analysed using the ideas from the hybrid theory. Given the evidence in the case, it will be studied which possible stories can be built that explain the explananda, how these stories can be supported, contradicted and compared and how the different evidential arguments attack each other. Before starting the analysis, my standpoint (Anderson et al., 2005, pp. 115–117) as an analyst should be clarified. An analyst's standpoint is dependent on four variables: time at which the analysis was performed, purpose with which the analysis was performed, materials available for analysis and the role of the analyst. In this case, the analysis was done post-trial, post-appeal and after the other members of "Gerede Twijfel" extensively reviewed and analysed the evidence. The current analysis is somewhat exceptional in that its main purpose to test the hybrid theory: by modelling the case in the theory

[1] Following (Crombag and Israëls, 2008), all names of persons other than Marjan van der E. are fictitious.

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_6, © Springer Science+Business Media B.V. 2011

presented in Chapters 4 and 5, it can be determined whether the theory is sufficiently expressive to correctly model a complex case in a natural way. However, this does not mean that the analysis serves other, more usual purpose, such as organization or evaluation of the evidence. Whilst the case has not been recreated in its entirety (some of the arguments and evidence have been left out of the analysis), the important evidence has been summed up in the table on p. 168 and the various different points of view and lines of reasoning have been rendered in their respective section in this chapter. Furthermore, the different accounts of what happened in the case are evaluated according to the criteria of the hybrid theory. Here the most important question is "what happened in the case" Consequently, my standpoint in this analysis can be characterized as that of a "historian" in that I do not regard the case from the viewpoint of either party (i.e. prosecution or defence) and I do not have the aim of either proving Marjan van der E.'s guilt or innocence. Furthermore, no assumptions are made as to the specific (investigative or decision-making context of the analysis). Finally, I will not concern myself with the legal aspects of the case.

As for the materials, the analysis in this chapter is mainly based on Crombag and Israëls' book and the Court of Appeals' judgement in the case. The information about the investigation (i.e. persons involved, testimonies and other evidence) is adapted from Crombag and Israëls' book and the judgement. The alternative stories are based on the judiciary's view of the case and the scenario's given in the book, but the exact way in which they are modelled and the method of analysis is unique to this book. Here it can be argued that the contents of the current analysis run the risk of being influenced too much by Crombag and Israëls' analysis and that for a proper analysis the original files should be studied. However, as was argued above the current study is not intended to provide new insights into the Anjum case, but as a case study for the hybrid theory as presented in this book.

## 6.1 The Murder of Leo de Jager

On the evening of the 24th of December 1997, the police find the lifeless body of Leo de Jager in the front yard of a boarding house in Anjum. The body is wrapped in a tent canvas and has a fractured skull. The police have acted on the statement of one Evert Beekman, who came to the police station around 19:00 that evening and pointed out the location of the body. The proprietor of the boarding house, Marjan van der E., and a friend and guest of hers, Marga Waanders, are taken into custody that night as the prime suspects in a possible murder case. Beekman also tells the police that there is another body to be found near the boarding house. The police start digging in a small field next to the boarding house and find the partially decomposed body of Herre Sturmans.

At the time of the case, the main suspect Marjan van der E. is in her fifties. She has studied biology and oceanography and, according to her brother, left almost all the jobs she held after her studies because of some quarrel. She also has a problematic relationship with her family. At some point in time, Marjan went to live in a small house in Moddergat, a hamlet about 5 km from Anjum. When she left Moddergat, she kept this house as a holiday home. In 1996, Marjan van der E. bought the boarding house in Anjum and became its proprietor.

The second suspect and important witness is Marga Waanders, a 40-year old woman who has regularly been visiting the boarding house since the summer of 1997. She has become friends with Marjan, who invites her over to Anjum for the Christmas holiday in 1997. Marga is supposed to take care of the boarding house because Marjan will visit Den Haag over the holidays.

The victim, Leo de Jager, is 26 years old at the time of his death. A poor student in school, he starts a job as a road worker at an early age. When he is 16, he starts drinking and regularly gets into trouble with the police. After a few failed relationships, he ends up in Moddergat where he rents Marjan's holiday home. Leo regularly visits the boarding house, where he is doing odd jobs for Marjan. Near the end of 1997, Leo befriends Pier, who also does odd jobs for Marjan at the boarding house. Through Pier, Leo hooks up with Bregje, Pier's ex-girlfriend. Bregje moves in with Leo in Moddergat and stays there until a day after his death.

The man who informs the police, Evert Beekman, is a 33-year old dealer in timber who lives in Anjum. At first he claims not to know Marjan really well. However, it later turns out that he has participated in a cannabis-growing operation in the barn of the boarding house together with Marjan and a friend of his, Jaap Kuilstra (see below).

The last person that needs to be introduced is Eef Tasman, a 52-year old man who regularly stays at the boarding house when is in the neighbourhood for his work. Tasman sometimes does administrative work for Marjan and might have been unwittingly involved in Marjan's attempts to hide the cannabis-growing operation.

The boarding house is situated on a through road, from which the front of the house can be clearly seen. There are two buildings on the site, the house and a barn, which is connected to the house. Seen from the road, there is a camping site left of this barn which is also part of the lot. A wide gravel drive leads up to the two front doors of the house and a dirt driveway leads to the barn and the camping site. In front of the house, there is a lawn which is partly surrounded by a hedge and partly by a low embankment.

The inside of the house (Fig. 6.1) can roughly be divided into two parts. The left part – consisting of a living room, a kitchen and a library/study – is the residential part of the building where Marjan lives. The right part, which has its own front door, contains the stairs to the first floor and is considered to be the guests' part of the building. The two parts are connected via a U-shaped hallway containing a built-in closet and a blue cupboard. At the time of Leo's death, Marga Waanders was staying in the house. Her bedroom has been marked in Fig. 46, which has been adapted from (Crombag and Israëls, 2008). "Leo's room" is the room where Marjan allegedly laid Leo to rest. There were no other guests at the time so bedroom 2 was empty (Marjan's bedroom was upstairs).

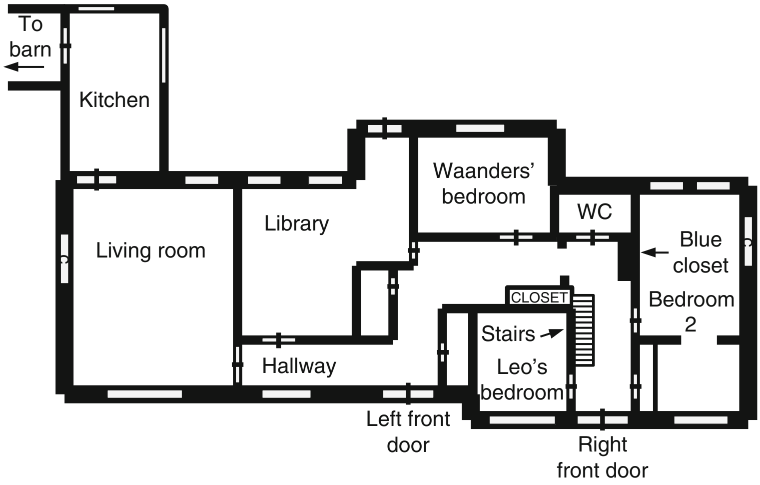



**Fig. 6.1** The interior of the boarding house

### *6.1.1 Before the 24th: The Cannabis-Growing Operation and Bank Fraud*

The 24th was not the first time that the police had come to the boarding house. A week before, on the 17th of December, the police had raided the house because they suspected that a large-scale cannabis-growing operation was taking place in the barn on the grounds of the boarding house. On the loft of the barn the police found about 2,100 cannabis plants and professional growing equipment. The estimated value of the plants was about 500,000 Dutch Guilders (around 224,000 Euros). At the time, Marjan said she rented out the barn in which the cannabis was found to a third party, a firm called Mandersman & Partners, and that she would show the police a lease contract. She then called Eef Tasman and asked him to draft a contract detailing the lease of the barn to Mandersman. Tasman, who claims he did not know the purpose of the contract, did as he was asked. Somewhere on the 24th of December Tasman went to the boarding house and gave Marjan the contract. The police investigation into the cannabis operation is still pending when Marjan is apprehended under suspicion of murder later that evening. The contract of which Tasman spoke is later found when the police search the house. Marjan denies any involvement in the operation.

During the investigation into Leo's death, more information about the cannabis-growing operation surfaces. Evert Beekman states that he and his friend Jaap Kuilstra had set up the operation together with Marjan. According to Beekman, Marjan already had plants on the loft of the barn together with two men from Leiden, but she wanted to get rid of these men because she did not trust them. After deceiving the two men by claiming the police had found the cannabis, Marjan proposed that Beekman and Kuilstra join her in the operation. Beekman claims that if the police would find out about the cannabis, Marjan would not mention Beekman and Kuilstra. After the police raided the barn on the 17th, Marjan allegedly told Beekman that she was going to use Leo as a front for the operation by making him sign the false contract Tasman was going to provide. The police found a form with Leo's signature and personal details in the house. The firm Mandersman & Partners really existed and was contacted, but no-one there knew of the cannabis operation.

Another interesting case the police came across when investigating the murder is a possible bank fraud. In January 1998 the police receive a phone call from an employee of the local bank in Anjum. A client named Veerman reported that more than 15,000 Guilders (more than 6,500 Euros) had been transferred from his account to several other accounts which he did not recognize. One of these accounts is Leo de Jager's and thus the police suspect that Leo's murder may be linked to this bank fraud. Further investigation shows that the money has been transferred using OLO's (optically readable transfer forms); normally such OLO's are only used internally but sometimes they are given to clients when the bank is out of normal forms. In the police's search of the boarding house, several OLO's are found. The police also find a bank statement of Veerman's, who lives in Moddergat near Marjan's holiday home and obviously the police now suspect Marjan of the bank fraud.

### 6.1.2 The Evidence in the Investigation into Leo's Death

This section contains an exposition of the most important pieces of evidence. First, the important witness testimonies will be discussed. During the investigation, the two most important witnesses were Evert Beekman and Marga Waanders, who both also have been suspects in the case. Both Beekman and Waanders gave extensive testimonies on which the judiciary based most of their story. The third person who might have known what happened to Leo, Marjan van der E., has always been treated as the main suspect. Her testimonies are less elaborate.

Below, the testimonies by Beekman and Waanders are summarized. In each case, first the testimonies in which the witnesses gave a general account of the events is given; this account is essentially a summary of the testimonies as recreated in the Court of Appeals judgement, combined with some further information from other testimonies.[2] After these main accounts, parts from other testimonies by the witnesses which are interesting for some reason will be discussed. Examples of such testimonies are testimonies that for some reason are inconsistent with the witness' main account, or multiple (vague) testimonies about some single event.

Note that here no arguments based on the evidence are built and that therefore no conclusions about, for example, witness credibility, will be drawn. The current section purely serves as an elaborate discussion of the elements in the set of evidential data $I_E$. After the discussion, the evidential data is summarized in Table 6.1.

---

[2] As was noted before, the information about other testimonies is taken from Crombag and Israëls (Crombag and Israëls, 2008)

#### 6.1.2.1 Beekman's Testimonies

In the investigation into Leo's death, Evert Beekman is an important witness, as it was Beekman who pointed out that there was a body in the first place. Beekman claims that Marjan told him she killed Leo when she came to Beekman's house to ask for help. Beekman himself is later also suspected of a crime, namely dragging Leo's body from the front door of the house to the lawn and concealing it under some frames of wire mesh.[3] Furthermore, Beekman was involved in the cannabis operation together with Marjan and Jaap Kuilstra (see above).

Beekman gives his first important testimony on the night of the 24th. Eleven days later, when he is suspected of helping Marjan move the body, he gives a second testimony, which differs from the first on some small points. The account below is mainly based on Beekman's testimony from the 5th of January 1998, which was summarized in the Court of Appeals judgement:

On the 23rd of December, around 18:30, Marjan arrived at Evert Beekman's house and told him that she had "bumped off" Leo. Marjan, whom Beekman is familiar with because he has sold her some timber in the past, asked Beekman to come to the boarding house to help her drag the body. Marjan then went back to the boarding house. Beekman told his girlfriend, Aaltje, what Marjan had said and that he would go to the boarding house to check it out. Ten minutes after Marjan left, Beekman walked to the boarding house, which is about a 5 min walk from his house.

When he arrived he was let in through the left front door by Marjan and immediately went left to the living room, where he met Marga Waanders. Marjan then told Marga that Leo had puked in the hallway and left the living room. Marga and Beekman stayed in the living room and talked about nothing in particular; Marga told Beekman her eyes were irritated so he advised her to put a wet washing cloth on them. Marga went into the hallway to get this and came back a short while later. All in all, Marga and Beekman were together in the living room for about 10 min.

Beekman then went into the hallway and saw Marjan scrubbing the floor in front of Waanders' bedroom door. Beekman saw blood on the floor of the hallway and a trail of blood the size of the back of a head leading through the hallway to the right front door. He went to the bathroom and when he came back, he noticed that Marjan had almost finished scrubbing. The trail of blood was not entirely clean but lighter than before. Marjan then locked the door to Leo's bedroom and said to Beekman she would tell Marga that Leo was sleeping there.

Beekman then went outside with Marjan. In front of the right door he saw the body of a man under a tent canvas and he recognized the man as Leo de Jager. Beekman noticed that Leo had no pulse and that Leo's skull was fractured and that there were six or seven "bald" spots on Leo's head. He saw Marjan pick up and throw away a stone. Marjan told him Leo could not stay in front of the door. Beekman and Marjan rolled Leo's body in the tent canvas and tied it up with a rope. Beekman then went home.

Around 2 o'clock at night, he returned and found Marjan in the kitchen. Beekman and Marjan dragged Leo's body to the lawn where they put it under some frames with wire mesh that were lying around. Beekman then went home again.

[3] Under Dutch law, the transportation or concealment of a body with the intent of hiding the fact or cause of death is considered a crime.

The next morning, around 11:00, Beekman went to Kuilstra and told him what happened. Kuilstra advised him to report the crime, because otherwise Beekman might be seen as an accomplice to Leo's murder. It is not until 19:00 that Beekman acts upon this advice and visits the police together with Kuilstra.

In his earlier testimony on the 24th Beekman withheld some of the above information. In this first testimony, Beekman states that Marjan said she bumped off "someone" instead of "Leo". Furthermore, according to this earlier testimony Beekman only identified Leo and went home. Only in his later testimony on the 5th of January does he admit he rolled the body into the canvas and went back in the night to drag the body to the lawn.

#### 6.1.2.2 Waanders' Testimonies

The second important witness is Marga Waanders, who was with Marjan when the police arrived at the boarding house the evening of the 24th. Waanders had arrived the previous day and was going to look after the boarding house when Marjan would go to Den Haag for the holidays. She was one of the last persons to have seen Leo alive and because of this she is also a main suspect early on in the case. Waanders is questioned multiple times and she gives several conflicting and different testimonies. The main testimony below is based on the Judgement of the Court of Appeals and is mainly an account of what happened before Beekman arrived around 19:00. For the most part, this account was not changed or contradicted by Waanders. Other events for which Waanders gave conflicting testimonies are recounted after this main testimony.

Waanders arrived at the boarding house between 13:00 and 13:30. Leo was in the living room with Marjan, where they were busy with some forms Waanders did not know the purpose of. Waanders noticed that Leo had bad motor functions, staring eyes and that he staggered when standing up. Waanders did not smell alcohol. Around 14:00 Waanders went out for groceries. She returned around 14:30 and noticed that Leo's reactions were slower than before she went out. Waanders then again left for town and returned around 16:30.

When Waanders returned, she met Marjan outside the house. Marjan was just about to cycle to town to buy something to calm Leo down. She also mentioned that Leo was in the barn; Waanders then found Leo in the back of the barn. She took him inside the house; while taking him to the living room, she had to support him. Waanders asked Leo if he used any medicines when she found him in the barn. Leo answered that normally he used around 23 medicines per day, but that he had not taken any in the preceding days or that day.

After a short while, Marjan came back and looked agitated at the fact that Leo was back in the house. Marjan got a glass of water in the kitchen; this water was probably warm as Waanders saw steam. Waanders saw that Marjan poured jenever in the glass of water. Marjan gave the glass to Leo, telling him it will calm him down, and he drank from it. The empty glass was again filled with warm water and jenever by Marjan and given to Leo. While Leo was drinking his second glass, Marjan took him to the hallway. She told Waanders that she was going to try to put Leo in bed.

Shortly after 17:15, Waanders started preparing dinner in the kitchen. While she was doing this, Marjan came in to fetch the bottle of jenever. Around 17:45 Waanders had finished preparing dinner, walked into the hallway and called for Marjan. Marjan replied with "I'll be right there". Around 18:10 Waanders again walked into the hallway to call Marjan for dinner. She saw Marjan and Leo in the hallway; Leo had collapsed on his knees. Marjan came up to Waanders and said something like "I've almost fixed it, I'll be right there". Waanders went back to the living room and heard someone walk on the gravel outside the house.

After a short while, Beekman and Marjan entered the room. Waanders smelt vomit in the hallway but she and Beekman were not allowed to see this by Marjan, who came to get a bucket from the kitchen to clean up the hallway. Waanders talked with Beekman and told him about her allergies and that her eyes were irritated. He advised her to put a wet washing cloth on her eyes and Waanders went out into the hallway to get a cloth from the closet opposite her room. After she came back in the living room, Marjan came in and took Beekman with her. When Waanders had finished her dinner, some time after Beekman's visit, she went to brush her teeth in her room.

That evening, Waanders and Marjan drank a few glasses of wine and went for a long walk with their dogs, around three quarters of an hour. Waanders then went to bed. The morning of the 24th she got up around 11:30; both Marjan and Leo were not there. When Marjan returns, she tells Waanders that Leo was picked up by Bregje earlier that morning.

Notice that the above testimony is vague about the events between 17:15 and 19:00, the supposed time that Marjan went to see Beekman, scrubbed the blood out of the hallway, dragged Leo outside to the front door and, with Beekman's help, dragged the body to the lawn. This is not because Waanders does not say anything about this time but because she gives multiple vague and sometimes conflicting testimonies about this period.

In one of her first testimonies, Waanders says that she had "images" of Marjan hitting Leo on the back of his head with something the size of a shoe in the hallway, some time between 17:45 and 18:15 (when Waanders called Marjan for dinner). A day later, however, Waanders says that the fact that she had these "images" does not mean she actually *saw* Marjan hit Leo. In fact, she then claims she did *not* see Marjan hit Leo.

Waanders is also asked by the police if she saw anyone lying in the hallway. First she claims she did not see anyone lying in the hallway. Later, however, she states that she saw Leo lying in the hallway.

Leo lay next to the blue cupboard under a blanket with his head towards the front door. He lay very still and looked pale and his hair was darker than what she remembered from earlier that afternoon.

In another testimony she even claims she took a few steps towards him. Waanders is unclear about the exact time at which she saw Leo lie in the hallway: first she says it was when she went to brush her teeth after dinner (Beekman had already left at that point) but later she said it was when she got the washing cloth from the closet (when Beekman was still in the boarding house).

Waanders says she saw something of Marjan's cleaning activities: she saw Marjan bustle about with a bucket of water, supposedly because Leo had vomited. When Waanders later went to the toilet, the floor of the toilet was wet (the water was not clear but also not bloody) and she dried it with toilet paper. First, Waanders says this was around midnight, when she and Marjan got back from walking the dogs. Later, Waanders states that this was at some point during Beekman's visit in the early evening.

Later, Waanders gives other interesting information about this visit to the toilet:

At some point during or just after dinner (Beekman had not yet left), Waanders went to the toilet. She was on the toilet for about 5 min when she heard a persistent bumping sound in the hallway. When she got off the toilet, she saw Leo's legs and shoes disappearing out of sight. Apparently, someone who Waanders could not see because he or she is round the corner of the hallway was dragging Leo. A short while later, she saw two shadowy figures standing outside through the front living room window. Given their rough shape, it is possible that these figures were Marjan and Beekman; they were talking and bended over to look at something on the ground.

One interesting thing to note is that Waanders never saw blood in the hallway. She did not see it in the early evening of the 23rd, nor did she see blood later that evening or on the 24th.

#### 6.1.2.3 Marjan's Testimonies

Almost right from the start, Marjan van der E. is the police's main suspect. The evening she is apprehended she says she is confused and does not give a testimony. During the investigation, Marjan is questioned more than 50 times; still she does not give a single, coherent statement of what exactly happened on the 23rd of December. Most of the time, she says nothing, or she claims that she feels threatened by Beekman. She denies that she has anything to do with the cannabis operation, Leo's drugged state before he died and Leo's death. Marjan's many brief testimonies about what happened on the 23rd can be compiled into the following summary:

Leo arrived in the morning. When Waanders arrived, Marjan and Leo were busy signing forms with Leo's personal details on them. Marjan later wanted to calm Leo down, so she went to town to get some alcohol around 16:30. When she got back, she gave Leo a grog, warm water with jenever to calm him down. At some point, Leo threw up in the hallway. Marjan suspected he was drunk and advised him to stay in the boarding house that evening.

At some point that evening, before 21:00, Marjan drove to Moddergat to Bregje and told her that Leo would stay in the boarding house that evening. Marjan said she would be back the following morning to take Bregje to an appointment at the Social Services in Dokkum (the nearest sizeable town at about 10 km from Anjum). Back at the boarding house Marjan went to check on Leo, who was fast asleep. Waanders went to bed and Marjan also went to bed not long afterwards.

Marjan only remembers that Leo threw up at a later stage of the investigation. At first, Marjan does not mention Beekman's supposed visit; later, she says that Beekman came to the house between 19:00 and 20:00 in the evening to fix the pen for the geese. In later testimonies she states that Beekman came to the boarding house that evening to give her the keys to the barn. The nightly visit around 2:00 about which Beekman spoke is not mentioned by Marjan and neither is the walk Waanders says she and Marjan took with their dogs. About the 24th of December, Marjan testifies the following:

Marjan got up around 9 in the morning. Leo had already left, probably on foot as his car was still parked outside the house. Marjan drove to Moddergat to pick up Bregje and then brought Bregje to Dokkum, first to Pier's house and later to the Social Services. Marjan then dropped off Bregje at Pier's house again and drove back to Anjum.

Together with Waanders she then went to Moddergat to get some papers and some of Bregje's clothing out of the holiday home. She locked the home with a new padlock because she said she did not want Bregje in the house anymore.

#### 6.1.2.4 Other Testimonies: Kuilstra, Bregje and Aaltje

In addition to the above testimonies of the main players in the case, the police also collect testimonies by others. Some of these testimonies yield new information; others only confirm or deny specific events from the above testimonies.

Kuilstra was the person who advised Beekman to go to the police and he accompanies Beekman the evening of the 24th. He stated that Beekman came to his house around 11:00 that morning and gave the above account of events. However, there are two slight differences between what Beekman tells the police and what Kuilstra says Beekman told him. Kuilstra states that, according to what Beekman told him, Leo was still in the hallway when Beekman arrived around 19:00 and that Marjan threw away a stone which had blood and hairs on it. Beekman told the police Leo was outside and he never mentioned the blood and hairs on the stone.

Bregje, Leo's girlfriend who also lived in the holiday home, confirms most of the events of the morning of the 24th that Marjan testified to:

Bregje asked Marjan how Leo was doing when she and Marjan were driving to Dokkum. Marjan avoided answering the question by asking if Bregje knew Leo's PIN code. Bregje did not know this, but together with Marjan she thinks of a possible code given the way in which Bregje had seen Leo's hands move across the keyboard of the cash machine. When they arrive in Dokkum, Marjan writes this code down. Pier, Bregje and Marjan then go to the Social Services, where it turns out Pier and Bregje do not have an appointment. Marjan buys a padlock in town and then drops Pier and Bregje at Pier's house.

Aaltje, Beekman's girlfriend, confirms that Marjan came to their house around 18:30 and that Beekman followed her to the boarding house some time later. She states that Beekman came home around 19:30 and said to her that Marjan had killed Leo by bashing his head in.

#### 6.1.2.5 Police Reports: Blood Markings and Temazepam

After the police find the body, the boarding house and its surroundings are searched for any further initial clues. The body lies on the lawn in front of the house, wrapped in a blue tent canvas and under another orange canvas and some wooden frames over which a wire mesh is stretched. There is also a trail with two grooves in the gravel from the right front door to the lawn. The police also find a heavy stone that could be the stone Beekman and Kuilstra spoke about.

Inside the boarding house a wad of paper with blood and hair is secured in a bin in Waanders' room. Furthermore, the police find the draft contract that Tasman made for Marjan which detailed the lease of the barn to Mandersman & Partners as well as a form with Leo's signature and personal details on it.

A day later officers of the forensics team come in to look for any traces and objects in the boarding house and the surrounding area. In the hallway of the house many traces of blood are found. In front of the door to Waanders' bedroom (see Fig. 6.1), a large bloodstain is found, covered by a rug. There are also bloodstains near the door to the WC, in front of the door to Bedroom 2 and inside near the right front door. Blood spatters are found on the walls, ground, doors and doorframes between Waanders' bedroom and the closet and left of the closet (from the perspective of Fig. 6.1). Many of these traces were still visible with the naked eye. The four stains are also visible on the wooden floor under the carpeting.

In the kitchen, the police find a strip of Temazepam, a sleep-inducing and muscle-relaxing drug, in a garbage bag. The strip contains ten empty capsules which have been cut open. In another garbage bag, three more empty strips are found. Furthermore, an empty can with a fresh orange peel and a large amount of empty Temazepam capsules is discovered. Finally, an empty medicine bottle is found; the label reads "Omeprazol" (a drug used in the treatment of stomach ulcers) and Leo's name. In the library, more Temazepam is found: 30 capsules in 2 boxes with the label "H. Sturmans" and a black pen case with strips containing a total of 63 capsules.

Later in the investigation, on the 5th of January, a small club hammer[4] is found in a bucket in the barn; both the hammer and the bucket contain watery bloodstains.

#### 6.1.2.6 Expert Reports: Autopsy, Toxicological Report and DNA Evidence

The autopsy of Leo's body shows that he has been killed by the use of violence. Leo had a fractured skull and significant brain damage, the combination of which can, according to the pathologist's report, unquestionably explain his death. The report further states that the fractured skull is the result of a strong, external and violent force, which can be caused by repeatedly being hit with an angular object and the brain damage is the result of a strong, external and violent force, which can be caused by being hit on the head or by falling.

A toxicological report is also made. This report states that Leo's blood contained 2.54 mg Temazepam per litre of blood and 0.46 mg of alcohol per litre of blood. The opinion of the toxicological expert is that the concentration of Temazepam in Leo's blood was higher than is to be expected in case of normal use of the medicine according to doctor's prescription. The toxicological expert also states that the concentration of Temazepam found in Leo's blood will cause dizziness and that it is likely that the combination of Temazepam and alcohol will have caused a state of impotence in Leo.

Finally, many (though not all) of the blood samples are sent to the laboratory for DNA analysis. It is the expert's opinion that the chances of the blood in the hallway and on Marjan's socks being from a random other person than Leo are much less than 1 in a million. The expert is less sure about the blood on the hammer: the chances of the blood on the hammer being from a random other person than Leo are 1 in 100 (hammer head) and 1 in 1,700 (hammer handle). A second opinion by

[4] A club or lump hammer is a is a heavy, one-handed hammer with a double-faced head.

another expert states that the chances of the blood on the hammer (head and handle) being from a random other person than Leo are less than 1 in 200 billion. The rest of the blood samples (i.e. on the cigarette, pager and tobacco) are not tested.

Table 6.1 below contains a summary of the most important pieces of evidential data. In the table, the data is categorized according to the part of the case it is relevant for and its type. For each piece of data a brief description of its contents is given. Furthermore, with each piece of evidence an identifier has been associated for ease of reference; in some cases this identifier is for the whole piece of evidence (e.g. search_financial) and in other cases, the individual contents of the data have been marked with a separate identifier (e.g. contract).

**Table 6.1** Evidential data in the Anjum case

| | |
|---|---|
| ***Police reports*** (***documents***)<br>*The cannabis operation* | |
| Police search (search_cannabis) | –2,100 cannabis plants and professional equipment<br>–Draft contract for the lease of the barn (contract)<br>–Form with Leo's details and autograph (form) |
| *The bank fraud* | |
| Police search (search_financial) | –Leo's bankcard in Marjan's purse<br>–OLO's and Veerman's bank statement in the boarding house |
| *Leo's murder* | |
| Police (forensics) search (search_Leo) | –Leo's lifeless body on the lawn in front of the house (body_Leo)<br>–Heavy, hand-sized stone in front of the house (stone)<br>–Trail in the gravel from front door to field (gravel_trail)<br>–Wad of paper with blood and hairs in Waanders' room (bloody_wad)<br>–Blood traces in the hallway (blood_hallway)<br>–Boxed capsules of Temazepam for Herre Sturmans in the library and loose capsules of Temazepam in the library. Empty strips of Temazepam in garbage (Tpam)<br>–Empty bottle of Omeprazol for Leo (Omeprazol)<br>–Small club hammer with blood on it in the barn (hammer) |
| ***Expert reports*** (***expert testimony***)<br>*Leo's murder* | |
| Leo's autopsy (autopsy) | –Leo had a fractured skull and significant brain damage<br>–The fractured skull is the result of a strong, external and violent force, which can be caused by repeatedly being hit with an angular object<br>–The brain damage is the result of a strong, external and violent force, which can be caused by being hit on the head or by falling<br>–The combination of the fractured skull and the brain damage can unquestionably explain Leo's death |

**Table 6.1** (continued)

| | |
|---|---|
| Toxicological report (tox_report) | –Leo's blood contained 2.54 mg Temazepam per litre of blood and 0.46 mg of alcohol per litre of blood<br>–The opinion of the toxicological expert that the concentration of Temazepam in Leo's blood was higher than was to be expected in case of normal use of the medicine according to doctor's prescription<br>–The opinion of the toxicological expert that the concentration of Temazepam in Leo's blood will cause dizziness and that it is likely that the combination of Temazepam and alcohol will have caused a state of impotence in Leo |
| DNA analysis 1 (DNA_1) | –An expert's opinion that the chances of the blood in the hallway and on Marjan's socks being from a random other person than Leo are much less than 1 in a million<br>–An expert's opinion that the chances of the blood on the hammer being from a random other person than Leo are 1 in 100 (hammer head) and 1 in 1,700 (hammer handle) |
| DNA analysis 2 (DNA_2) | –An expert's opinion that the chances of the blood on the hammer (head and handle) being from a random other person than Leo are less than 1 in 200 billion |
| ***Witness reports* (*witness testimony*)** | |
| *The cannabis operation* | |
| Marjan | –Marjan stated she rented out the barn to Mandersman & Partners and that she would show the police a lease contract (MvdE_Mandersman)<br>–Marjan was not involved in the cannabis operation (MvdE_deny_cannabis) |
| Beekman (Beekman_cannabis) | Marjan was in the cannabis operation together with Beekman and Kuilstra. Marjan said Beekman and Kuilstra's names would not be mentioned in case of a police raid. After the raid, Marjan told Beekman she was going to use Leo as a front |
| Tasman (Tasman_contract) | –Marjan asked him to draft a contract of lease for the barn<br>–Tasman made the contract and gave it to Marjan |
| Mandersman (Mandersman_contract) | –Mandersman & Partners states that there never was a contract with Marjan |
| *The bank fraud* | |
| Bank employee (Bank_employee) | –15,000 guilders were transferred from Mr. Veerman to 5 accounts, one of which was Leo's account<br>–The transfer was commissioned using special forms |
| Bregje (Bregje) | Marjan asked for Leo's PIN number |
| *Leo's murder* | |
| Beekman's earlier testimony (EB_early) | –Marjan came to Beekman's house and told him she killed someone<br>–Beekman did not say he helped drag Leo's body |
| Beekman's main testimony (EB_main) | Testimony about how Beekman was asked for help by Marjan and helped Marjan wrap the body in the canvas (see p. 167) |
| Marga Waanders' main testimony (MW_main) | Testimony by Waanders about Leo's drugged state and Beekman's visit (see p. 171) |

**Table 6.1** (continued)

| | |
|---|---|
| Waanders' images of Marjan hitting Leo (MW_image) | –Waanders had "images" of Marjan hitting Leo on the back of his head in the hallway |
| Waanders about Marjan hitting Leo (MW_not_hit) | –Waanders did not see Marjan hit Leo on the back of his head in the hallway |
| Waanders about Leo in the hallway 1 (MW_hallway_1) | –Leo lay next to the blue cupboard under a blanket with his head towards the front door<br>–This was when Waanders got the washing cloth from the closet in the hallway |
| (MW_hallway_2) | –Same as MW_hallway_1 only the time was when Marjan went to brush her teeth |
| Waanders' toilet visit (MW_toilet) | –At some point during or just after dinner (Beekman had not yet left), Waanders went to the toilet<br>–Waanders heard a persistent bumping sound in the hallway. When she got off the toilet, she saw Leo's legs and shoes disappearing out of sight. Apparently, someone was dragging Leo outside |
| Waanders about blood (MW_blood) | –When asked, Waanders consistently answered that she did not see blood in the hallway |
| Marjan's main testimony (MvdE_main) | Marjan's testimony about Leo's arrival and the trip to Moddergat (see p. 173) |
| Marjan's denials | –Marjan denies she gave Leo Temazepam (MvdE_deny_Tpam)<br>–Marjan denies she killed Leo (MvdE_deny_kill) |
| Kuilstra's testimony (Kuilstra) | –Beekman said Marjan threw away a stone with blood and hair on it |
| Aaltje's testimony (Aaltje) | –Marjan came to their house around 18:30 and Beekman went to the boarding house about 10 min later<br>–Beekman came back around 19:30 and told Aaltje Marjan had killed Leo |

## 6.2 An Analysis of the Case: Constructing Stories and Arguments

Now that the evidential data in the case has been summarized, the process of proof can start. Note that in an actual investigation often not all of the evidential data is available at the start of the process: often first hypothetical stories are constructed and then new evidence is sought to confirm or deny the story. Because in this analysis I take the standpoint of a historian it is assumed that no new information will be uncovered.

The method of analysis is based on the formal dialogue system from Section 5.5: the operations performed correspond to the speech acts in the dialogue game and, like in the protocol, only speech acts that are relevant to the case at hand are made. The relevant operations on a story or a set of evidence are linked to the critical questions and pitfalls mentioned in Section 4.4, the principles of which were subsequently formalized as the formal criteria in Section 5.4. In the case of alternative explanations for some event, the explanations will be compared and weighed according to the formal criteria.

The presentation of the stories and arguments in the case study will be "semi-formal": logical formulas are paraphrased as natural language sentences. Large stories that are intended to provide an overview of the events are best presented in natural language as a list of events. Note that in this presentation of stories, the simple causal links between the events are simply assumed even though they are not explicitly rendered. Individual small stories or parts of a larger story can be presented as a causal graph; this presentation is also convenient if the causal or abstraction relations in a story are being discussed. For the presentation of arguments, graph-like figures will also be used. Finally, attacks between multiple arguments will in some cases also be rendered as an attack graph (Fig. 3.7, p. 42). As the figures in Chapter 5 have shown, a translation of these figures into the formal theory is straightforward. In many cases, a figure will slightly abstract from the contents of the actual formal speech act that is moved in that not all events, arguments and causal relations that are moved with the relevant speech act are shown in the figure. It is indicated in the text when this is the case.

In order to improve the readability, the dialogue moves will not be rendered precisely in the way they would be moved in an actual formal dialogue (as it is done in the example in Section 5.6). For example, the players and speech acts are implicitly assumed. Furthermore, the discussion does not follow the back-and-forth format of an actual dialogue, where each explanation for the explananda is slightly improved until the turn switches to another player. In this analysis, explanations and their merits and deficiencies will be considered in turn, that is, first one explanation for a pivotal event and its supporting and contradicting arguments will be discussed and only then does the discussion switch to the next, alternative explanation.

In the analysis, an explanation will often be *extended* by, for example, explaining initial causes, predicting or explaining new evidence or refining causal relations. Furthermore, for clarity I will often first focus on a single part of a larger explanation and how such a part can be improved or worsened and only then *combine* the various parts into one larger explanation. The operations on explanations "extend" and "combine" have no formal equivalent in the dialogue game: every time an explanation is extended or explanations are combined, this is done in the game by a single *explain* move that contains the extended or the combined explanation (cf. condition 6a of Definition 5.5.13 and the examples in Section 5.6). However, for obvious reasons of readability and space, I will not always render the complete explanation but simply say that "an explanation is extended" or "explanations are combined" and implicitly assume that each time the full explanation is given in a single *explain* move.

In the original case, almost all of the evidence is of the documentary type. Hence, the *document* prima facie reason (reason 5 in Definition 5.2.1) should officially always be used to infer, for example, the testimony and only then can the prima facie reason associated with the evidence type mentioned in Table 6.1 be used. In the current case, however, this first inference step based on the *document* prima facie reason will in the case of a testimony not be modelled and only the testimonial inference will be modelled. This is in my opinion not a dangerous shortcut, as at no point in the case was the authenticity of the documents in question.

As was already noted, only the part of the case concerning Leo's murder will be discussed. First, the reasoning surrounding Leo's cause of death, the murder weapon, the location of Leo's death and Leo's drugged state before he died will be discussed (Section 6.3). While in the original case this was not a matter of much debate, it allows for the illustration of some of the basic concepts of the theory such as causal and evidential reasoning, attacking arguments, supporting events and causal relations. Furthermore, the two alternative explanations for the murder weapon and their corresponding evidence are relatively easy to keep track of so they are good candidates for a detailed illustration of the dialogue game and the hybrid theory.

In Section 6.4, the main story in the case, namely the view held by the judiciary that Marjan killed Leo and that she worked alone, will be discussed. Based on the judgement of the Court of Appeals and Crombag and Israëls' book, the view of the judiciary will first be reconstructed and subsequently improved upon by elaborating on Marjan's involvement in Leo's drugged state (Section 6.4.1), on Marjan's possible motives for killing Leo (Section 6.4.2) and finally on the events after Leo died (Section 6.4.3); Section 6.4.4 briefly summarizes the improved judiciary's story that was constructed in the previous subsections of 6.4. The discussion in Section 6.4 focuses less on the specifics of the formal dialogue game but rather presents a general procedural approach to the analysis of evidence and hypothesis with the hybrid theory and the formal criteria.

In Section 6.5, alternatives to the judiciary's story will be considered. There are two issues in the case for which there are interesting and important alternatives, namely what or who caused Leo to be in a drugged state and who killed Leo. In Section 6.4 the judiciary's story that Marjan drugged Leo is discussed; Section 6.5.1 discusses the alternative, that Leo took the Temazepam himself. Similarly, Section 6.5 discusses the alternatives to the story from Section 6.4 that Marjan killed Leo. These alternatives all involve Beekman in various roles. Section 6.6 briefly compares the alternatives from Sections 6.4 and 6.5. Section 6.7 ends the case study by briefly evaluating its implications for the hybrid theory and the dialogue game.

## 6.3 Cause of Death, Murder Weapon and Leo's State

In this section, it will be shown how, given the evidence, it can be established what the cause of Leo's death was and what the murder weapon was. Note that in the actual case as well as in Crombag and Israëls' book, this was not really a matter of debate. The discussion illustrates specific elements of the hybrid theory on reasoning with evidence.

The main explanandum in the case is Leo's death: that Leo was dead can be justifiably inferred from the police report of the search of the boarding house (body_Leo) and the pathologist's report (autopsy) with the *document* and the *expert* prima facie reasons, respectively. First it should be established how he died: could it have been an accident or was he killed? The pathologist's report (autopsy) gives some clues as to what happened: Leo's death was caused by the combination of a fractured skull and brain damage. Both these direct causes of death could themselves have been caused by being hit on the head with an angular object; the brain damage could also have been caused by falling. Because we are clearly dealing with causal information here ("The fractured skull [. . .] which can be *caused* by repeatedly being hit [. . .]"), it makes sense to model the conclusions of autopsy as a causal explanation, or rather, as two possible causal explanations (Fig. 6.2).

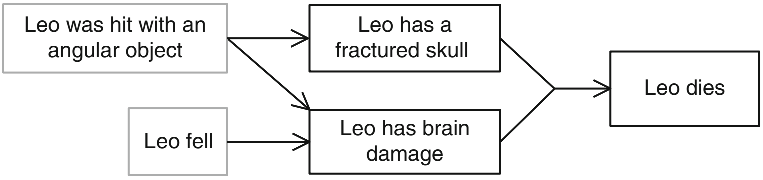



**Fig. 6.2** Explanations for Leo's death

In Fig. 6.2, all the events and causal generalizations except for Leo was hit with an angular object and Leo fell are directly supported by autopsy. The events that Leo died, that he has a fractured skull and that he has brain damage were literally stated by the expert and can hence be inferred with the *expert* prima facie reason. Both the brain damage and the fractured skull are the result of "strong, external and violent force" and this force can be caused by being hit in the head (in both cases) or by falling (in the case of the brain damage). Note that the expert does not say whether Leo fell or whether he was hit in the head: he only supplies possible causes in his testimony. The events Leo was hit with an angular object and Leo fell are therefore evidential gaps, which is shown by rendering the boxes for the two non-supported events in a lighter shade. Now, there are two explanations for the event Leo has brain damage: he could have fallen or he could have been hit on the head. The explanation Leo fell, however, does not explain the explanandum because it does not explain the explanandum Leo dies; this is because Leo has a fractured skull is needed to explain Leo dies (the generalization reads Leo has a fractured skull $\wedge$ Leo has brain damage $\Rightarrow_C$ Leo dies). The fractured skull is not explained by Leo falling so we can safely assume that Leo was hit with an angular object.

### *6.3.1 The Murder Weapon*

In the case, there were two possible angular objects: the hammer (hammer) and the stone (stone). Leo could have been hit on the head with either of these. Reasoning about the identity of objects is done using abstraction generalizations (see Section 5.3.3). In this case, the fact that there are two possible objects with which Leo was hit is modelled as there being two different explanations for the

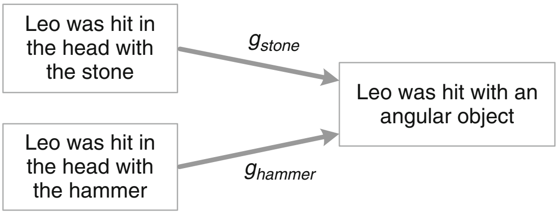



**Fig. 6.3** Two possible murder weapons

event Leo was hit with an angular object. These explanations then involve abstraction relations rather than causal relations, viz. Figure 6.3.

These generalizations $g_{stone}$ and $g_{hammer}$ basically say that club hammers and stones are angular objects. Based on these generalizations, there are now two explanations for Leo was hit with an angular object: $S_{stone}$ = {Leo was hit with the stone, $g_{stone}$} and $S_{hammer}$ = {Leo was hit with the hammer, $g_{hammer}$}.[5] Combined with the explanation from Fig. 6.2, both $S_{stone}$ and $S_{hammer}$ explain the explanandum Leo dies.[6] It should now be determined which one of the two explanations for Leo was hit with an angular object is the most likely by starting a dialogue game with two players, where one defends $S_{stone}$ and the other $S_{hammer}$. For simplicity, the exact move-by-move dialogue game will not be written out but I will first analyse $S_{stone}$ using the possible dialogue moves and then expand and analyse $S_{hammer}$. After this, the two explanations will be compared.

The idea that a stone might have been used to kill Leo essentially came from Kuilstra (see evidence Kuilstra). Beekman said that Marjan threw away a stone but he did not say anything about blood. Kuilstra, on the other hand, stated that he heard from Beekman that the stone had blood and hair on it (Fig. 6.4 on p. 181).

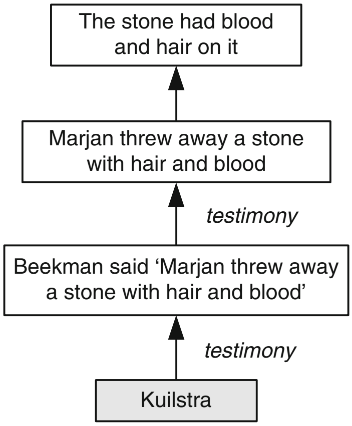
Argument $A_{Kuilstra}$[7] is an example of hearsay evidence: the witness testimony reason is applied twice in succession. Hence not only Kuilstra's but also Beekman's veracity, objectivity and observational sensitivity can be questioned. In order to support the explanation $S_{stone}$ with this argument, the explanation should be expanded so that it incorporates the event the stone had blood and hair on it, which is supported by Kuilstra's testimony. Now, the event Leo was hit with the stone in $S_{stone}$ together with the causal generalization $x$ was hit with the stone $\Rightarrow_C$ the stone had blood and hair on it ("hitting someone in the head with a stone may cause blood and hair to be on the stone") can be used to *predict* that there may be blood on

[5] Explanations will generally be named as $S_{name}$, where *name* is the specific name identifying that explanation.

[6] In the rest this section, the explanation [Leo was hit with an angular object, Leo was hit with an angular object $\Rightarrow_C$ Leo has a fractured skull $\wedge$ Leo has brain damage, Leo has a fractured skull $\wedge$ Leo has brain damage $\Rightarrow_C$ Leo dies] from Fig. 6.2 is implicitly assumed. Thus, the explanandum Leo dies follows from any explanation that explains Leo was hit with an angular object.

[7] Arguments will generally be named as $A_{name}$, where *name* is the name identifying the argument.

**Fig. 6.4** Argument $A_{Kuilstra}$ for the blood and hair on the stone



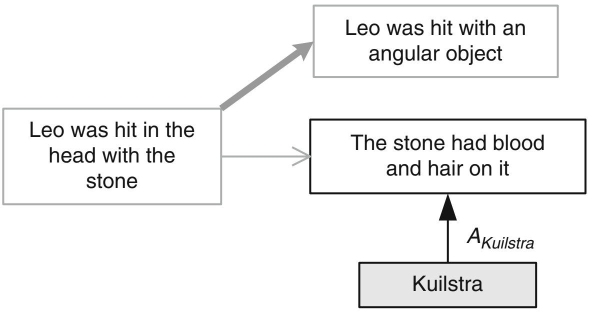




**Fig. 6.5** $S_{stone}$ supported by Kuilstra's testimony

the stone.[8] The new explanation is now directly supported (see Definition 5.4.2) by Kuilstra's testimony, viz. Fig. 6.5. This explanation explains that Leo was hit with an angular object (and thus that Leo died, Fig. 6.2 on p. 179) and the conclusion of Kuilstra's testimony that the stone had blood and hair on it. Note that even though this observation does not directly explain Leo was hit with an angular object, it adds to the quality of $S_{stone}$ because the updated $S_{stone}$ is now supported by Kuilstra.[9] This shows how new events (story-consequences) can be predicted from events in the story and that, if these new events are supported by evidence, such a prediction can improve the story (because its evidential support is improved).

[8]Recall from the end of Section 5.6 that prediction can be done with the *explain* speech act: *explain* ({Leo was hit with an angular object} [Leo was hit with the stone, Leo was hit with the stone $\Rightarrow_A$ Leo was hit with an angular object, Leo was hit with the stone $\Rightarrow_C$ the stone had blood and hair on it]).

[9]Formally, the unsupported $S_{stone} = \{$Leo was hit with the stone, $g_{stone}\}$ is still in the set of explanations *Expl* and there are now two versions of $S_{stone}$. However, for clarity only the newest version (which is supported by Kuilstra) is shown.

This prediction of the stone had blood and hair on it can also have a negative effect on $S_{stone}$. The police found a clean stone, similar to the one described by Beekman and Kuilstra, with no hair and blood on the premises (stone). Furthermore, no bloody stone was found. It can therefore be argued that the evidence stone contradicts the stone had blood and hair on it: the only stone similar to the one described did *not* have blood and hair on it. However, this may be explained by assuming that someone wiped off the stone after Leo was hit with it. This assumption is not entirely implausible, as a wad of tissue with blood and hair was found (wad). By predicting that the stone must have been clean after someone wiped it off, a new version of $S_{stone}$ can be given (Fig. 6.6) so that it is compatible with the evidence stone and wad.

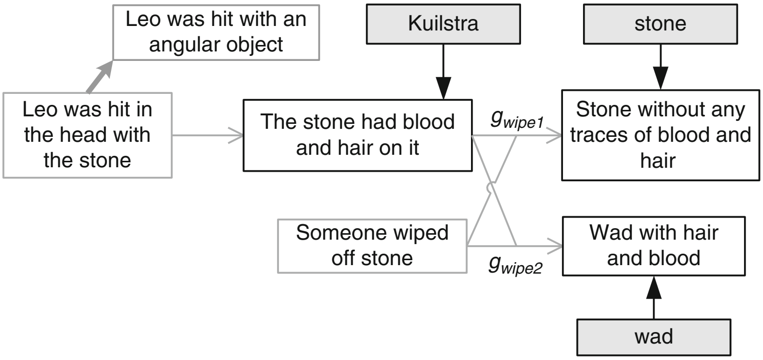



**Fig. 6.6** The expanded $S_{stone}$

There are two causal relations that can be expressed with a generalization that has the conjunction the stone had blood and hair on it **and** someone wiped off the stone in the antecedent: $g_{wipe1}$ ("wiping off a stone with hair and blood on it with a tissue will cause all visible traces of blood and hair to be removed from the stone") and $g_{wipe2}$ ("wiping off a stone with hair and blood on it with a tissue will cause there to be hair and blood on the wad of tissue"). The new version of $S_{stone}$ shown here is supported by three pieces of evidence: Kuilstra, stone and wad. Here, the evidence stone and wad are documents and the facts that there were a stone and a wad at the crime scene are inferred through the *document* reason.

There are two weaknesses in the explanation $S_{stone}$ as rendered in Fig. 6.6. The first weakness concerns the evidential part of the theory, namely the hearsay argument $A_{Kuilstra}$. Beekman never claimed there was hair and blood on the stone, so it seems that either Kuilstra lied or was mistaken about what Beekman told him, or Beekman did not tell the police what he told Kuilstra (that there was blood on the stone). Beekman has no direct reason to withhold this information from the police: it does not in any way harm his position if Marjan threw away a stone with blood. Thus, we can assume that the inference from Kuilstra's testimony in $A_{Kuilstra}$ is undercut (Fig. 6.7) because either Kuilstra misheard what Beekman told him (undercutter on Kuilstra's observational sensitivity), Kuilstra misremembers what

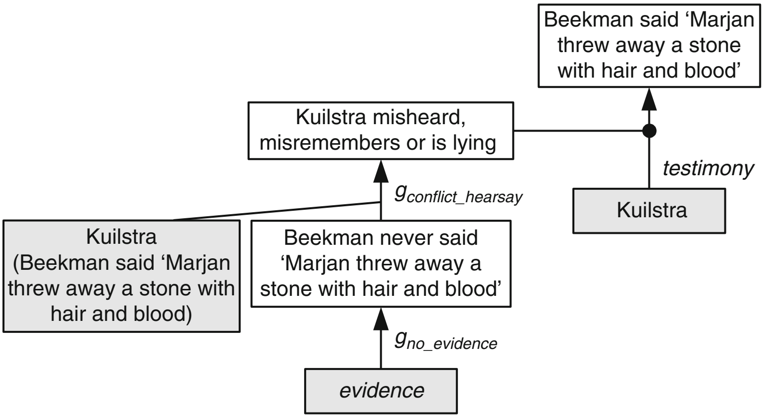



**Fig. 6.7** Undercutting $A_{Kuilstra}$

Beekman told him (undercutter on Kuilstra's objectivity) or Kuilstra is deliberately lying to incriminate Marjan and thus protect his friend Beekman (undercutter on Kuilstra's veracity).

The generalization $g_{conflict\_hearsay}$ expresses that if one witness (here: Kuilstra) testifies that another witness (Beekman) said something and this other witness never said it, then the first witness (Kuilstra) must be mistaken or lying. Note that this inference also has Kuilstra's witness testimony as a premise. The other generalization $g_{no\_evidence}$ expresses that if there is no evidence for a proposition then this proposition is probably not true.[10] In this way, the negation of propositions which are not in the total collection of evidence (*evidence*) can be inferred. This generalization should be used carefully: it presupposes that Beekman was asked about the stone or that he at least had the opportunity to tell about the stone and that Beekman himself did not lie or misremember. Would any of these things be the case, we would have an exception to $g_{no\_evidence}$ and $A_{Kuilstra}$ would in effect be reinstated. For the moment, however, it is assumed that $A_{Kuilstra}$ is defeated and that thus the testimony Kuilstra no longer supports $S_{stone}$.[11]

Another weakness in $S_{stone}$ is that it contains an explanation which is implausible, namely $g_{wipe1}$ ("wiping off a stone with hair and blood on it with a tissue will cause all visible traces of blood and hair to be removed from the stone"). Recall that we can argue for the implausibility of a causal generalization in a story by contradicting

[10] This is similar to the logic programming principle of *negation as failure*, according to which failure to derive $p$ can be used to derive $\neg p$.

[11] According to Definition 5.4.2, if an argument based on evidence is overruled (i.e. decisively defeated) then the evidence does not support its conclusion.

it with an argument based on general knowledge (Definition 5.4.8).[12] In this case, we have to construct an argument for the invalidity of the generalization $g_{wipe1}$ (i.e. an argument with $\neg(valid(g_{wipe1}))$ as its conclusion). This argument can be based on general knowledge, i.e. "it is general knowledge that wiping off a stone with a tissue will usually not remove all visible traces of blood". With the *general knowledge* prima facie reason, the reason for the invalidity of the generalization can then be inferred. The attack on the story $S_{stone}$ can be rendered as in Fig. 6.8. At this point, the evidential support (Definition 5.4.3) of $S_{stone}$ is wad and its causal implausibility is $g_{wipe1}$ (Definition 5.4.8).

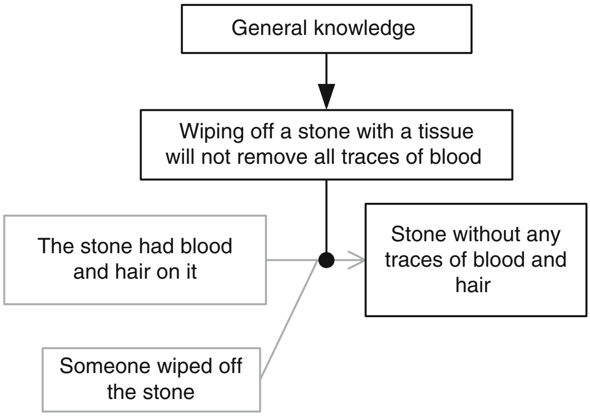



**Fig. 6.8** Argument against the plausibility of a generalization

Now let us consider $S_{hammer}$. An argument can be constructed for the conclusion that the hammer had Leo's blood on it (Fig. 6.9). An interesting question here would be how to exactly calculate the chances that the blood on the hammer was Leo's. This involves not only a simple probabilistic calculation, but also has to do with the accrual of reasons: do we give equal weight to both analyses or do we trust one expert more than the other? Do we consider the fact that the two estimates are clearly different? These issues were briefly considered in Section 3.1.2; for now, it will simply be assumed that the chance that the blood was someone else's is significantly small enough.

The fact that there is a club hammer with Leo blood on it may mean that Leo was hit with this hammer. Here, an analogous line of causal reasoning as was used for $S_{stone}$ can be applied to $S_{hammer}$: it may be possible that someone wiped some of the blood and the hair off the hammer after Leo was hit with it. In this way, both the hammer with Leo's blood on it and the bloody wad of tissue can be explained by Leo was hit on the head with the hammer (Fig. 6.10). The causal generalizations

[12]If we were to contradict the generalization with evidence we would not be arguing for the implausibility, which is established independently from the evidence, but rather for the evidential contradiction (Definition 5.4.5).

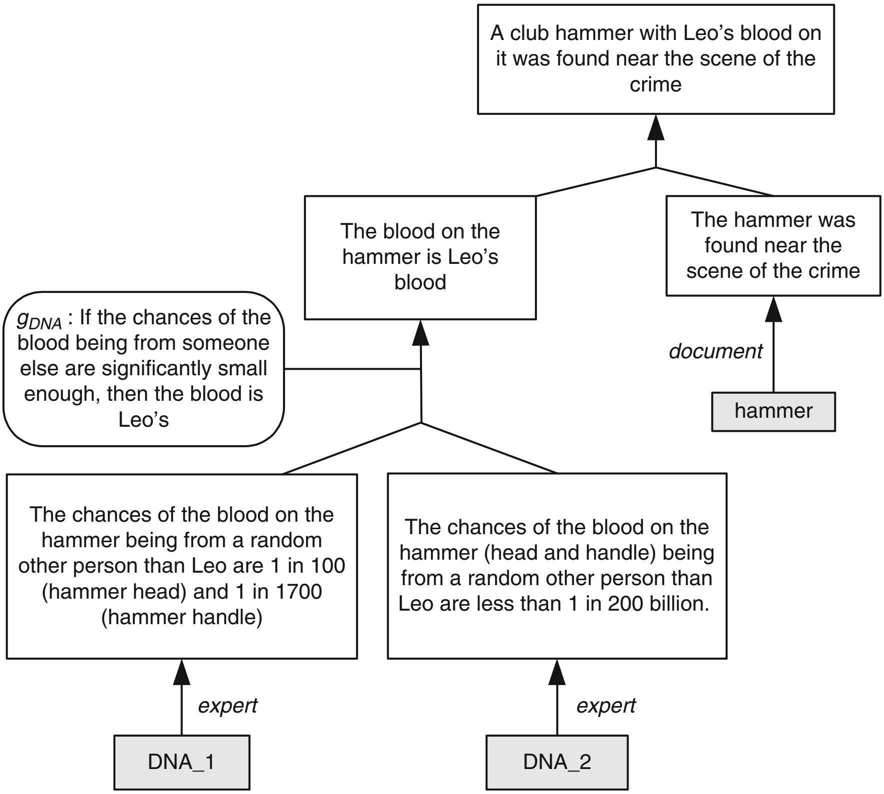



**Fig. 6.9** Argument $A_{hammer}$ about the blood on the hammer

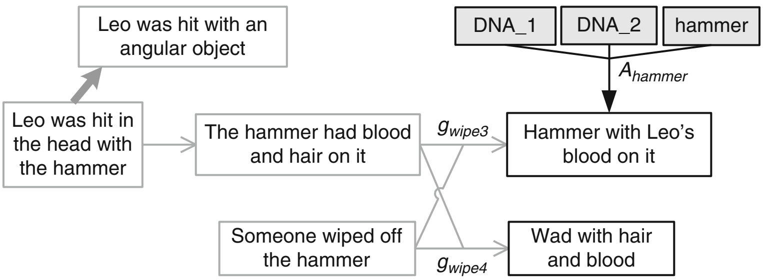



**Fig. 6.10** The new version of $S_{hammer}$

$g_{wipe3}$ ("wiping off a hammer with hair and blood on it with a tissue will cause all the hair and some of the blood to be removed from the hammer") and $g_{wipe4}$ ("wiping off a hammer with hair and blood on it with a tissue will cause there to be hair and blood on the wad of tissue") are similar to $g_{wipe1}$ and $g_{wipe2}$; the important difference (apart from the object being wiped off) is that unlike $g_{wipe1}$, $g_{wipe3}$ does not argue that all traces of blood will be removed. Because of this, $g_{wipe3}$ cannot be contradicted with an argument similar to the one in Fig. 6.8 (p. 184) and $g_{wipe3}$ can be regarded as more plausible than $g_{wipe1}$.

The explanation $S_{hammer}$ from Fig. 6.10 and $S_{stone}$ from Fig. 6.6 (p. 182) are very similar: both explain that Leo was hit with an angular object through an abstraction link and both also explain other observations that follow from evidential data. However, the evidential support of $S_{hammer}$ (DNA_1, DNA_2, wad, hammer) is higher than that of $S_{stone}$ (wad, stone). It may be argued that DNA_1 and DNA_2 should count as one piece of evidential data, because they represent essentially two corroborating reasons for the same conclusion. However, $S_{hammer}$ then still has a higher (justified) evidential support. In addition to this, the plausibility of $S_{stone}$ is worse than that of $S_{hammer}$ so either way, $S_{hammer}$ is the better explanation (according to Definition 5.5.9). This conforms to my intuitions on the two stories: $S_{stone}$ is a story based on hearsay evidence about some bloody stone that was never actually found and $S_{hammer}$ is a story about the one actual angular object that was found and that has Leo's blood on it.

The above discussion concerns detailed explanations about the cause and manner of Leo's death. It is shown how the various criteria for explanations (i.e. evidential support, causal plausibility) can be used to attack and defend explanations. A precise dialectical dialogue as proposed in Section 5.5 is not shown. However, it can be seen how the basic principles of the dialogue guide the above process of reasoning. Assume that there are two adversaries, one who thinks that Leo was killed with the stone and one who thinks Leo was killed with the hammer. The first player starts by proposing and supporting $S_{stone}$. This explanation is then attacked by the other player in various ways: a causal link is deemed implausible and a supporting argument is undercut. However, this is not enough: from Definition 5.5.10 (current winner) it follows that $S_{stone}$, however implausible or contradicted it might be, is still the best explanation simply because it is the only one. A second explanation $S_{hammer}$ should therefore be constructed to provide an alternative. This explanation should also be supported by evidential data so that it is ranked higher than $S_{stone}$. Notice that here the importance of the evidential data over the plausibility of the story plays a role: even though $S_{stone}$ is less coherent than $S_{hammer}$, as long as $S_{stone}$ is supported by evidence and $S_{hammer}$ is not, $S_{stone}$ is still the best explanation. Supporting $S_{hammer}$ with the DNA evidence makes it the best current explanation. If the discussion were to continue, the player supporting $S_{stone}$ would now have to make his explanation better or somehow contradict $S_{hammer}$. Given the available evidence, however, it does not seem that he has many options. Hence, it is for the moment safe to assume that Leo dies because he was hit on the head with the hammer.

### *6.3.2 The Location Where Leo Died*

Before the current analysis turns to the question who hit Leo and why, first a few other findings from the police search and the DNA evidence will be modelled. $S_{hammer}$ can be further extended by making the location of the murder explicit and incorporating some of the events that happened after Leo was hit. From the evidence it follows that Leo's blood spatters and Leo's blood stains were found in the hallway (Fig. 6.11). These two arguments, which use the same generalization $g_{DNA}$ as the argument $A_{hammer}$ (Fig. 6.9, p. 185), allow us to infer not only that Leo was hit with the hammer, but also that he was hit in the hallway and that he lay in the hallway after he was hit.

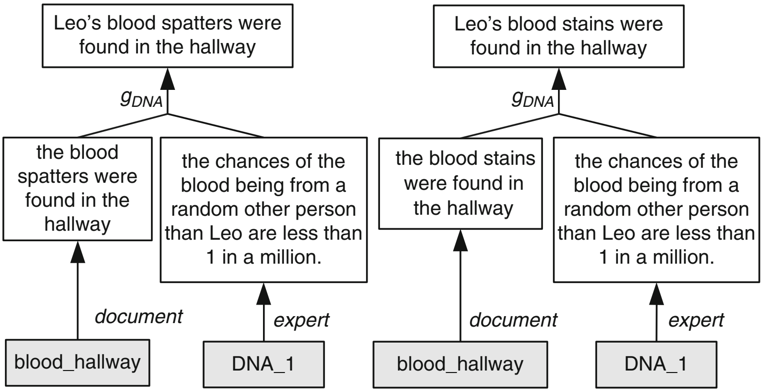



**Fig. 6.11** Arguments $A_{splatters}$ and $A_{stains}$ about the blood found in the hallway

The event Leo was hit in the head with the hammer in the hallway is a more detailed version of Leo was hit in the head with the hammer in the same way as of Leo was hit in the head with the hammer is a more detailed version of Leo was hit with an angular object, viz. Figure 6.12. Through this chain of abstraction generalizations, any explanation that includes Leo was hit in the head with the hammer in the hallway explains Leo was hit with an angular object (and therefore also Leo dies). Now, if we assume the story $S_{hammer}$ from Fig. 6.10, any explanation that includes Leo was hit in the head with the hammer in the hallway explains the observations wad with hair and blood and hammer with Leo's blood (because Leo was hit with an angular object explains these observations). Say that we have the explanation in Fig. 6.13 for the observations that Leo's blood spatters were found in the hallway and Leo's blood stains were found in the hallway. If we now accept that the newest version of $S_{hammer}$ is the combination of the explanations in Fig. 6.10 (for the bloody hammer, p. 185), Fig. 6.12 (for the specifics of Leo being hit) and Fig. 6.13 (p. 188), then this new $S_{hammer}$ is a story about how Leo was hit in the hallway with the hammer and how this may have caused the blood on the hammer and the wad and the blood

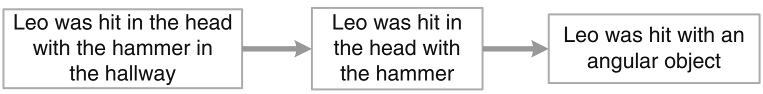



**Fig. 6.12** Reasoning from specific to general events with abstraction generalizations

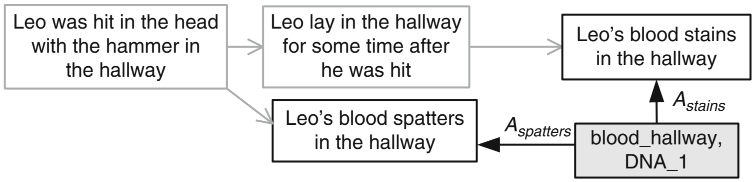



**Fig. 6.13** Explaining the conclusions of the blood evidence from the hallway

in the hallway. This explanation is then the best version of $S_{hammer}$ thus far and it is supported by (DNA_1, DNA_2, wad, hammer, blood_hallway).[13]

### 6.3.3 Leo's Drugged State

Before the stories about *who* killed Leo are discussed first another important issue, namely Leo's drugged state before he died, will be discussed. According to the toxicological report (tox_report), the concentration of Temazepam in Leo's blood was significantly higher than it would be with normal use of the medicine (Fig. 6.14).

The toxicological report also states that "the concentration of Temazepam found in Leo's blood will cause dizziness", that Leo's blood furthermore contained a slight percentage of alcohol and that "it is likely that the concentration of Temazepam and alcohol will have caused a state of impotence in Leo". Regarding Leo's state of impotence and his dizziness, Waanders mentioned multiple times in her testimony (MW_main) that Leo was stumbling and that he had to be supported and that he barely responded to her or Marjan. From this it can be more or less directly inferred that Leo was both dizzy and in a state of impotence. Waanders also said she saw Marjan give Leo a combination of Jenever and hot water ("a grog"), something which Marjan confirmed. The causal reasoning surrounding Leo's state and its supporting evidence can be modelled as in Fig. 6.15. That Marjan gave Leo the grog

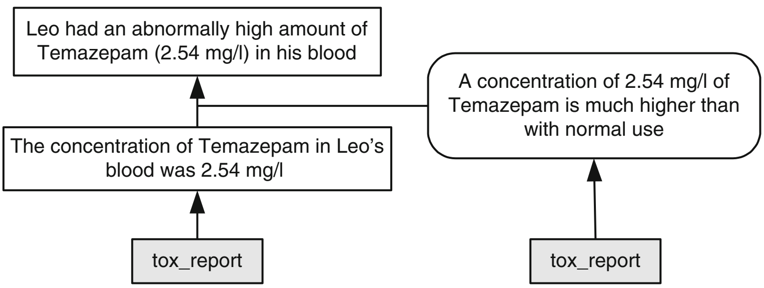



**Fig. 6.14** Argument $A_{Tpam}$ about the amount of Temazepam in Leo's blood

[13]Note that in the formal dialogue game, this newest version of $S_{hammer}$ (i.e. the combination of Figs. 6.10, 6.12 and 6.13) will have to be moved in one single *explain* move.

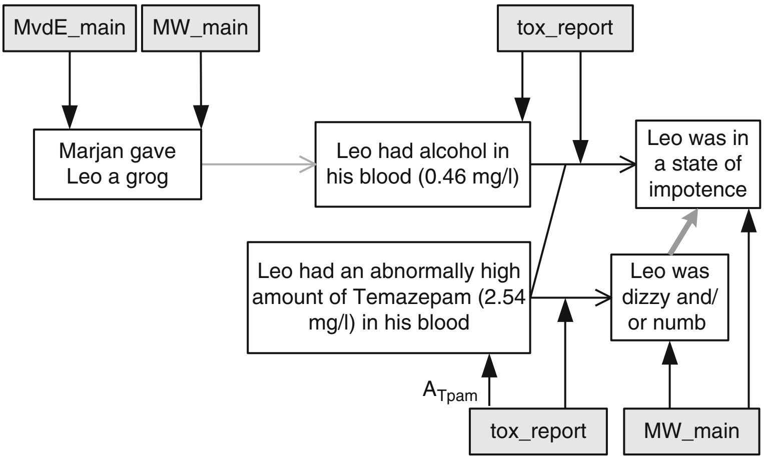



**Fig. 6.15** Leo was under the influence of alcohol and Temazepam

most likely caused him to have alcohol in his blood; together with the Temazepam this caused a state of impotence. The Temazepam alone would have caused dizziness and this state of dizziness can be seen as a state of impotence (hence the abstraction link).

## *6.3.4 Summary*

In the above sections it was shown that in all probability, Leo was hit with a hammer in the hallway and that before his death, he ingested some alcohol and a high amount of Temazepam. The two short stories $S_{hammer}$ (the combination of Figs. 6.10, 6.12 and 6.13) and that Leo was under influence of alcohol and Temazepam (Fig. 6.15) will in the rest of this analysis be considered proven. In both of the cases, there was no doubt as to whether the expert reports (autopsy and tox_report) were credible. In the case of the murder weapon, $S_{hammer}$ can be considered the best explanation given the evidence and the only reasonable alternative, $S_{stone}$, has been adequately tested and compared to $S_{hammer}$. Regarding the alcohol, both Marjan and Waanders stated that Marjan gave grog to Leo so we can reasonably assume this is the case; it is hard to think of another reason Marjan would admit to doing this. The complete story thus far can be rendered as in Fig. 6.16. For readability purposes, Fig. 6.16 slightly "summarizes" some of the causal reasoning. For example, the chain of generalizations from Leo was hit on the head to Hammer with Leo's blood has been shortened and so has the chain from Leo was hit on the head to Leo dies. Furthermore, the evidence that supports this explanation (MvdE_main, MW_main, tox_report, DNA_1, DNA_2, wad, hammer, blood_hallway, autopsy) has not been explicitly rendered (the directly supported states, events and causal relations have, however, been rendered in a darker shade than the unsupported states, events and relations). In the rest of

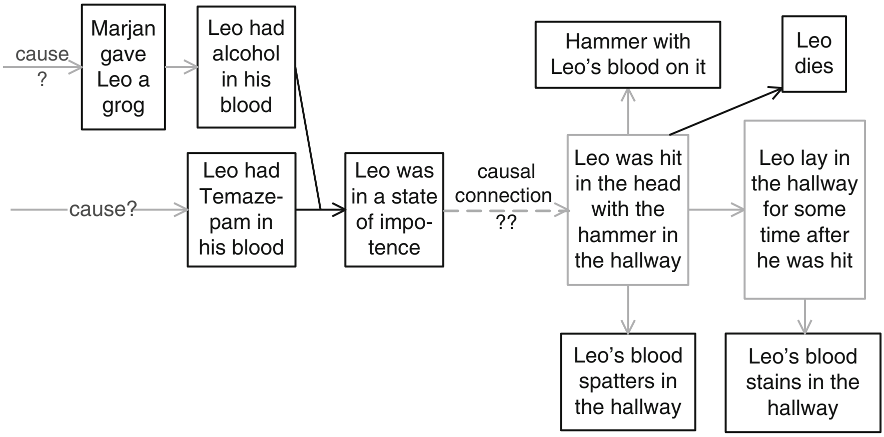



**Fig. 6.16** The events proven thus far

this analysis, the best version of the explanation $S_{hammer}$ and the explanation about Leo's state will be considered proven.

## *6.3.5 Initial Evaluation*

Thus far, both causal and evidential reasoning were needed in the case. For example, the reasoning about the cause of death, what caused the blood spatters to be in the hallway and what caused Leo to be in a drugged state is analyzed with causal relations; while these parts can perhaps also be modelled in an evidential argument, this would involve changing all the causal relations to evidential generalizations, which is gives rise to problems. For example: "a wad with hair and blood on it is evidence for the fact that someone wiped the hair and blood off a stone" (see Section 2.3.2). This can be problematic, as such an "inversion" of the direction of reasoning is only allowed when the cause usually causes the effect; in this case, we would not say that bloody wads are usually caused by wiping off a stone. Another disadvantage of a purely evidential approach is that the experts explicitly mentioned causal generalizations in their reports and changing these into evidential generalizations would require an interpretation step which brings us further away from the meaning of the generalization as originally intended by the expert.

Evidential reasoning allows the individual pieces of evidence such as expert reports, testimonies and other documents to directly support the events in the story. Furthermore, the argument $A_{Kuilstra}$ seems to be an intuitive way of talking about hearsay evidence; it allows for a detailed attack on the credibility of one of the witnesses. The (relatively simple) evidential arguments in the case focus the analysis by first allowing us to infer clear and interesting observations (i.e. the bloody stone, Leo's blood spatters, Leo's blood on the hammer) which can then be explained with a story. For example, in this case we are mainly interested in explaining *why*

*Leo had such a high concentration of Temazepam in his blood*. That he had a concentration of 2.54 mg/l of Temazepam in his blood and that this can be seen as a high concentration is certainly important (see Fig. 6.14, p. 188) but not an interesting observation for which we want to give alternative explanations. In other words, the evidential arguments allow us to focus on interpreting the evidence so that we have clear states and events that can fit into one or more understandable stories.

So far, the hybrid theory seems to provide all the tools necessary for the analysis. The procedural way of providing alternatives and trying to prove or disprove these alternatives works well in the discussion about the murder weapon. Modelling alternative "identities" of a person or object (in this case the angular object with which Leo was killed) with abstraction links provides a natural way of performing such reasoning by explicitly visualizing it. The formal criteria for the quality of a story provide adequate guidelines for the analysis by showing which parts of explanations can be improved or worsened. In particular, the fact that the notion of plausibility is independent from the evidence accurately separates the discussion about a story's conformity to the evidence and a story's inherent coherence. Finally, the ordering of explanations $S_{stone}$ and $S_{hammer}$ according to the criteria complies with our intuitions about why $S_{hammer}$ is the better story.

The above sections aimed at providing a detailed example of the dialogue game and the hybrid theory. In the further analysis of the case, the dialogue game and the criteria for comparing explanations will be assumed implicitly and it will not be constantly discussed in detail what the evidential support and plausibility of the explanations are. However, the case study is still intended as a test of the hybrid theory. The procedural treatment of the explanations in the case, where an explanation is proposed in a basic form and then expanded, supported and further refined, will be continued. Furthermore, the operations on the explanation are always aimed at either improving or worsening the explanation according to the criteria. Section 6.7 discusses the implications of the case study for the hybrid theory and the dialogue game in more general terms.

From Fig. 6.16, a few important issues that need to be resolved arise. First, why and by whom was Leo given the abnormally high dosage of Temazepam? Second, why and by whom was Leo hit on the head? And third, are these two events (i.e. the Temazepam and Leo's death) somehow linked? The course of events *after* Leo died might also be important but only to determine the identity of the perpetrator (and possibly to determine if any other crimes were committed, see footnote 3 in this chapter).

As the Temazepam and alcohol were both found in Leo's stomach, it can be safely assumed that the Temazepam and alcohol were orally ingested by Leo at a time when he was still alive. The question is of course if Leo was aware of this, if someone else gave him the Temazepam secretly or if he was somehow coerced into taking the Temazepam. From the three main testimonies (Beekman's (EB_main), Waanders' (MW_main) and Marjan (MvdE_main)) it can be gathered that Marjan, Waanders' and Leo himself were the only ones that had the physical opportunity to give Leo the Temazepam. In this analysis, it is assumed that the fact that

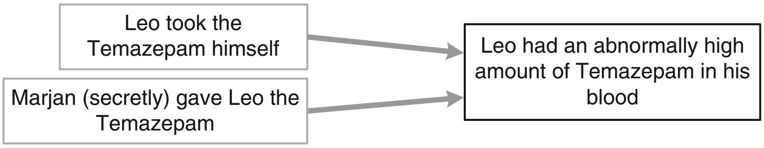



**Fig. 6.17** Identity of the person who drugged Leo

Waanders had no motive whatsoever to either drug or kill Leo exonerates her from any suspicion.[14] The possible explanations for Leo's drugged state are then as in Fig. 6.17. Notice that thus it is possible to talk about the identity of someone with abstraction links: each suspect has his or her own explanation and the best one should be chosen.

We know that Leo died because he was hit on the head with the hammer. This is not a complete story, as it provides only an *action* and a *consequence*. The identity of the person who hit Leo and this person's goals and motivations have not yet been determined. The evidence points to four possible persons who had the opportunity to kill Leo: Leo himself, Marjan, Waanders and Beekman. The possible suicide hypothesis (i.e. that Leo killed himself) can be discarded because the idea that Leo committed suicide by repeatedly hitting himself in the head with a hammer is quite implausible. That Marjan, Waanders and Beekman had the physical opportunity to hit Leo can be directly inferred from EB_main, MW_main and MvdE_main. Again, Waanders will not be considered here because she had no motive. The two possible explanations that Leo was hit on the head are as in Fig. 6.18.

The two explanations in Fig. 6.18 can be connected to the explanandum Leo dies in the way described in the previous sections. In the following sections, the alternatives that explain Leo's state of impotence and his death will be discussed. In Section 6.4, the judiciary's view that Marjan drugged and killed Leo will be discussed; Section 6.5.1 discusses the possibility that Leo took the Temazepam himself and 6.5.2 discusses various scenarios in which Beekman is the killer.

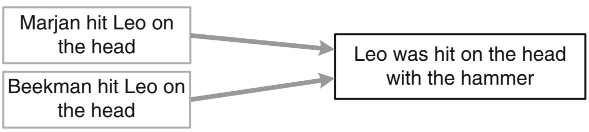



**Fig. 6.18** Identity of the killer

[14]Whilst it is possible to have a discussion about Waanders' possible motives in the hybrid theory, this will not be elaborated upon in this analysis.

## 6.4 The Judiciary's View: Marjan Drugged and Killed Leo

Let us start with the story the Court accepted, namely that it was Marjan who gave Leo the Temazepam so she could hit him over the head with the hammer in the hallway of the boarding house. In the judgement by the Court of Appeals, a single explanation or clear story is not given by the Court; it has to be guessed from the contents of the testimonies the Court cites as reasons for its conviction what the Court thinks happened on the 23rd of December. The Court mainly accepts the course of events as testified by Waanders in MW_main and Beekman in EB_main, the combination of which leads to roughly the following story:

*23rd of December*

| | |
|---|---|
| before 13:00 | Leo arrived at the boarding house (MvdE_main). |
| 11:00–13:00 | Marjan (secretly) gave Leo a significant amount of Temazepam. (*evidential gap*) |
| 13:00–16:30 | Waanders arrived at the boarding house (MW_main).<br>The Temazepam had its effect on Leo, making him dizzy and numb. (MW_main, tox_report)<br>Leo signed the meaningless forms (MW_main, MvdE_main, forms). |
| 16:30–16:45 | Marjan went to town to get alcohol (MW_main, MvdE_main).<br>Leo was in the barn for no reason, Waanders took him to the house. (MW_main) |
| 16:45–17:45 | Marjan gave Leo several glasses of grog (MW_main, MvdE_main).<br>Marjan took Leo to the hallway, "to put him in bed" (MW_main).<br>Waanders started preparing dinner in the kitchen (MW_main). |
| 17:45–18:10 | Waanders called Marjan several times, Marjan is busy (MW_main).<br>Waanders saw Marjan and Leo on his knees near the blue closet in the hallway (MW_main).<br>Waanders started dinner in the living room (MW_main). |
| 18:10–18:30 | Marjan hit Leo on the head with a hammer in the hallway (location indirectly supported by DNA_1, blood_hallway through $S_{hammer}$, see Fig. 6.13, p. 188; action itself is an *evidential gap*).<br>Leo died because he was hit on the head with the hammer. (indirectly supported by DNA_1, DNA_2, hammer, autopsy through $S_{hammer}$, Fig. 6.10, p. 185)<br>Leo lay bleeding on the floor of the hallway for sometime after he is hit (indirectly supported by DNA_1, blood_hallway through $S_{hammer}$, see Fig. 6.13).<br>Marjan dragged Leo outside |
| 18:30 | Marjan arrived at Beekman's house, said to Beekman that she has "bumped off Leo" and asked him to come to the boarding house to help her get rid of Leo (EB_main, Aaltje). |

| | |
|---|---|
| 18:45> | Beekman arrived at the boarding house and entered through the left door, he was taken to the living room by Marjan (MW_main, EB_main).<br>Beekman and Waanders talked while Waanders was eating her dinner, Marjan left for the hallway (MW_main, EB_main).<br>Waanders smelled vomit; Marjan was fussing about with a bucket because "Leo puked" (MW_main, EB_main).<br>Waanders went into the hallway to get a wet washcloth from the closet (MW_main, EB_main). |
| ± 19:00–19:30 | Beekman went to the hallway (MW_main, EB_main).<br>Marjan was still scrubbing the floor of the hallway (EB_main).<br>There was blood on the floor of the hallway (EB_main, blood_hallway).<br>There was a trail of blood towards the front door (EB_main).<br>Beekman and Marjan went outside, where Leo's dead body lies near the right front door under a tent canvas (EB_main).<br>Beekman and Marjan wrapped Leo's dead body in the tent canvas outside near the front door (EB_main). |
| ± 19:20 | Beekman went home (EB_main, Aaltje). |
| 19:20–20:00 | Waanders went to brush her teeth in her room |
| 1:30–2:00 | Beekman went back to the boarding house (EB_main).<br>Beekman and Marjan dragged Leo's body to the front garden. (EB_main, body_Leo) |

This story will be named $S_{Marjan\_Jud}$. Notice that behind each event the evidential data that supports the event is mentioned. Most of the events are supported directly by the data through its associated prima facie reason (e.g. from EB_main it can be concluded with the *testimony* reason that "Beekman goes home", see the various prima facie reasons in Definition 5.2.1). Other events, most of which were discussed in Sections 6.3.1, 6.3.2 and 6.3.3 as part of $S_{hammer}$ or the causal reasoning about Leo's state of impotence, are supported indirectly in that they support the best explanation on, for example, the murder weapon (i.e. $S_{hammer}$). The reader is referred to the above sections for a discussion of these events. Furthermore, as was already noted in footnote 83, formally any new explanation has to be moved in its entirety (see Definition 5.5.6 and condition 6 of Definition 5.5.13). However, for readability I will zoom in on certain aspects of the case and implicitly assume, for example, the explanations as given in the preceding sections.

In its judgement, the Court only mentions the evidential data and does not provide the above story. Notice that there are three important evidential gaps which are filled with assumed events, namely "Marjan gave Leo the Temazepam", "Marjan killed Leo" and, perhaps somewhat less important, "Marjan dragged Leo outside". These evidential gaps concern the identity of the person who committed the act. For example, there is evidence for the fact that Leo died from the effects of being hit on the head with the hammer. However, this does not support the fact that it was *Marjan* who hit him on the head. When there is no evidence for some events, these events have to be inferred from the other circumstances for which there is evidence. One way of creating plausible circumstances for an event involves "fitting" the event in a plausible story that is supported by evidential data. In this way, the story, or the circumstances, makes it more plausible that the events in question happened.[15] In the current case, the most important events (namely that it was Marjan who drugged and killed Leo) are not directly supported by evidence, so the reader is left guessing as to what really happened and why. The above story is a first attempt to fit the events which are not directly supported into a coherent story. In the next few sections, this coherence will be analysed and improved by assuming some new events and relations.

### 6.4.1 Marjan's Motives for Drugging Leo

An important evidential gap in the judiciary's story is the identity of the person(s) who gave Leo the Temazepam. We are looking for a coherent explanation of the following explananda: Leo had an abnormally high concentration of Temazepam in his blood, Leo had alcohol in his blood and Leo was in a state of impotence. In Fig. 6.15 (p. 189), the alcohol in Leo's blood is explained by assuming that Marjan gave Leo a grog. This explanation is obviously not complete, as it only includes an action (Marjan gave Leo the grog) and a consequence (Leo had alcohol in his blood); no goals or motivations for Marjan are given. The Temazepam in Leo's blood is not further explained at all. The Court seems to think that Marjan gave Leo the Temazepam; in order to make this plausible, a complete explanation that explains why Marjan gave Leo the alcohol and the Temazepam must be given. In its judgement, the Court argues that "Marjan gave Leo the Temazepam to bring him in a state which facilitated her killing him". In other words, Marjan wanted to bring Leo in a state of impotence so that she could kill him (Fig. 6.19).

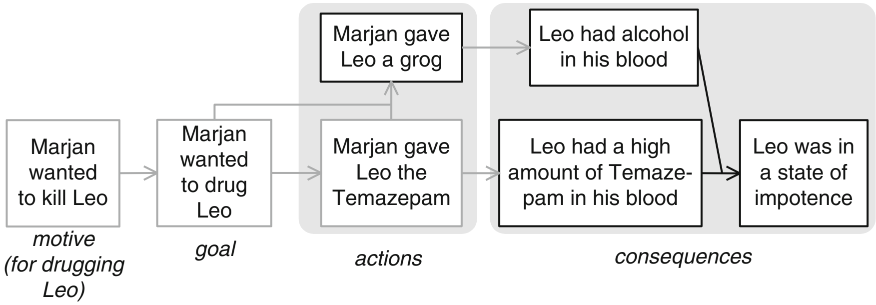



**Fig. 6.19** Marjan's reason to drug Leo ($S_{M\_Tpam}$)

[15]This is the "gap-filling" function of stories that has been mentioned in Section 4.2, where gaps in the evidence (*evidential gaps*) are filled with events that fit the story.

Recall that the gray rounded boxes mean that all the events in the box correspond to the respective component of the story scheme (see Section 5.3.4). In the above explanation, Marjan's immediate goal is to bring Leo in a state of impotence; this is motivated by her will to kill him. In order to reach this goal, she first gives him Temazepam and after that some alcohol in the form of a grog. According to the toxicological report, this caused Leo to be in a state of impotence (Fig. 6.15). The explanation completes the intentional action scheme: the observed consequences are clearly caused by actions, the actions are performed to further a goal and this goal is motivated. The way in which the events Fig. 6.19 are supported by evidential data was already discussed earlier (see Fig. 6.15).

The above explanation is open to several attacks against its plausibility. First, it can be argued that the generalization on which the first causal link is based – person $x$ wants to kill another person $y \Rightarrow_C x$ wants to bring $y$ in a state of impotence – is not complete; $x$ will only want to render $y$ helpless if $x$ is incapable of killing $y$ without first bringing $y$ in the state of impotence. The generalization can be refined by assuming that Marjan drugged Leo because she wanted to kill him *and* because Leo was stronger than Marjan. The second generalization that can be further refined is Marjan wanted to bring Leo in a state of impotence $\Rightarrow_C$ Marjan gave Leo the Temazepam.[16] This implicitly assumes that Marjan had the Temazepam on hand: without access to Temazepam, Marjan could not have given it to Leo. This implicit assumption should therefore also be made explicit. Figure 6.20 shows the two refined generalizations.[17]

The complete explanation $S_{M\_Tpam}$ consists of the combination of the explanations in Figs. 6.19 and 6.20. $S_{M\_Tpam}$ is fairly coherent, as it seems to complete the intentional action scheme and the causal generalizations in it can be reasonably accepted. Note that by refining the causal generalizations as in Fig. 6.20, the number of evidential gaps increases and that refining generalizations in this way might point to directions where new evidence can be found. The explanation in Fig. 6.20 is supported by the arguments given in Fig. 6.21. In this figure, $g_{stronger}$ stands for "26-year old men are generally stronger than 51-year old women" and $g_{secret}$ stands for the generalization "Hidden Temazepam capsules out of which the powder has been

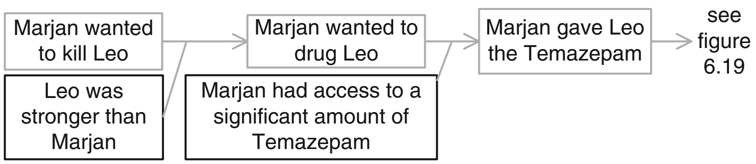



**Fig. 6.20** Refining the generalizations in $S_{M\_Tpam}$ (Fig. 6.19)

[16]Here the generalization is rendered in its case-specific form (cf. Section 2.1.3).

[17]Note that the present dialogue game does not include a "refine" speech act and that hence a new explanation has to be given which incorporates the refined generalization.

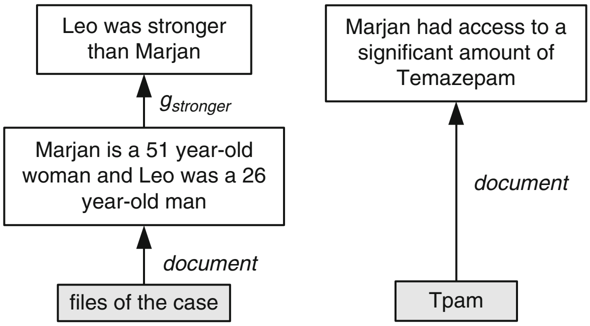



**Fig. 6.21** Arguments supporting $S_{M\_Tpam}$

taken may point to the fact that someone administered the Temazepam secretly". Notice that the argument based on empty_Tpam improves the evidential support of $S_{M\_Tpam}$ but does not directly support Marjan secretly gave Leo the Temazepam. The explanation $S_{M\_Tpam}$ is now reasonably well-supported. However, because the main action Marjan gave Leo the Temazepam is not directly supported we also need to consider alternative explanations for the fact that Leo had such a high concentration of Temazepam in his blood. Note that $S_{M\_Tpam}$ is a sub episode of the larger explanation that Marjan killed Leo as Marjan wanted to bring Leo in a state of impotence is a subgoal of the higher goal Marjan wanted to kill Leo. Hence we also have to show that the goal Marjan wanted to kill Leo fits into a plausible story about intentional actions; this will be discussed in the next section.

One alternative explanation assumes that Marjan gave Leo the Temazepam but with a different motive. Recall from Section 6.1.1 that Marjan was the suspect in the case about the cannabis operation and that Marjan needed a contract of hire for the barn in which the cannabis plants were found. Furthermore, Beekman stated that Marjan told him she was going to use Leo as a front. This possibly meant that she needed Leo's signature on the false contract. Figure 6.22 shows a small story about

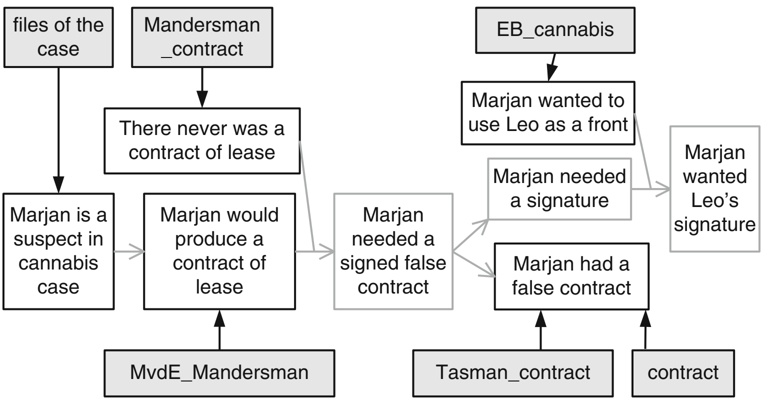



**Fig. 6.22** Story about Marjan's need for a false contract

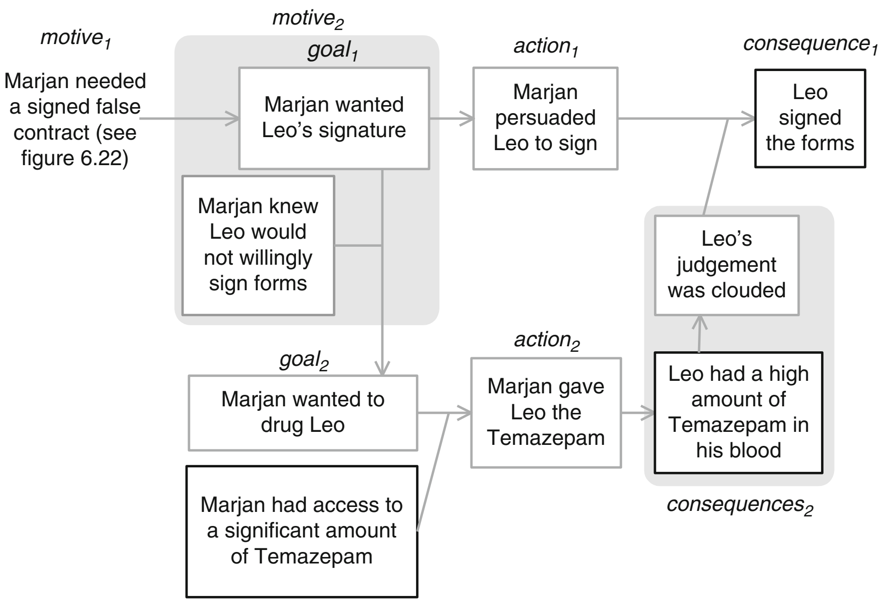



**Fig. 6.23** Marjan wanted Leo to sign the forms ($S_{M\_forms}$)

why Marjan would want Leo's signature; this story is supported by the evidential data in the cannabis case.

The story in Fig. 6.22 can act as a motive for why Marjan wanted Leo to sign the forms: she needed his signature on the false contract. Now, there is a possibility that maybe Marjan gave Leo the Temazepam to cloud his judgement so that he signed the forms which she could use for this contract (Fig. 6.23). The explanation in Fig. 6.23 not only explains why Leo had such a high amount of Temazepam in his blood but also why Leo signed the apparently meaningless forms that were found. Notice that the explanation consists of a main episode and a sub-episode (see Fig. 3.21 on p. 65). The goal of the main episode (*goal1*) together with the fact that Leo would probably not sign the forms when he was clear-headed acts as a motive for Marjan's subgoal to cloud Leo's judgement with the Temazepam. The arguments for the conclusions that Marjan had access to Temazepam and that Leo had Temazepam in his blood (not explicitly shown in Fig. 6.23) were given in Fig. 6.21 and Fig. 6.15, respectively. The causal link between the concentration of Temazepam and the state Leo's judgement was clouded is not supported by evidential data: while it is known that Temazepam can also cloud someone's judgement and make a person slightly reckless, no expert explicitly testified this (the expert only testified that the Temazepam might have brought Leo in a state of impotence). Finally, the event that Leo signed the forms is supported by MW_main, MvdE_main and the fact that forms were found with Leo's signature (forms).

There is another alternative or perhaps additional motive for Marjan to drug Leo which has to do with the bank fraud. Recall that someone managed to transfer a significant amount of money to Leo's account from Veerman's account. Quite possibly, it was Marjan who authorized this transfer with her OLO's and Marjan was most likely involved in the bank fraud (Fig. 6.24). So Marjan would need Leo's bankcard and PIN number in order to withdraw the money from Leo's account. Leo's bankcard was found in Marjan's purse the following day and Marjan explicitly asked Bregje for his PIN number. This can point to the fact that Marjan had the motive shown in Fig. 6.25 for drugging Leo. This explanation is similar to $S_{M\_forms}$: Marjan wants something of Leo and suspects (quite rightly) that Leo will not give this when he is sober so she drugs him with the Temazepam. Now there are essentially three explanations for Leo's drugged state: Marjan drugged him because she wanted to kill him ($S_{M\_Tpam}$, Fig. 6.19 and Fig. 6.20), Marjan drugged him because she wanted his signature ($S_{M\_forms}$, Fig. 6.23) and Marjan drugged him because she wanted his PIN number and bankcard ($S_{M\_fin}$, Fig. 6.25).

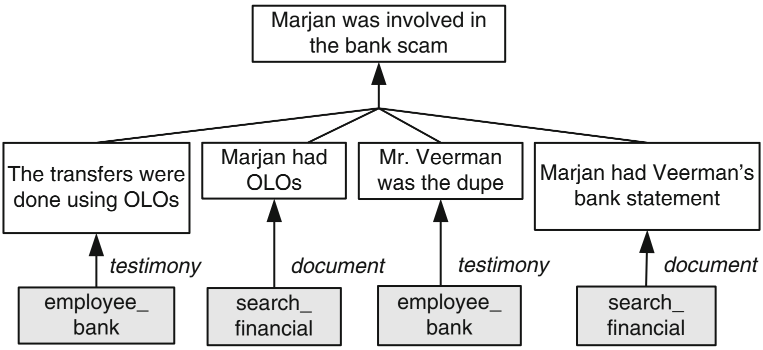



**Fig. 6.24** Argument $A_{fraud}$ for Marjan's involvement in the bank fraud

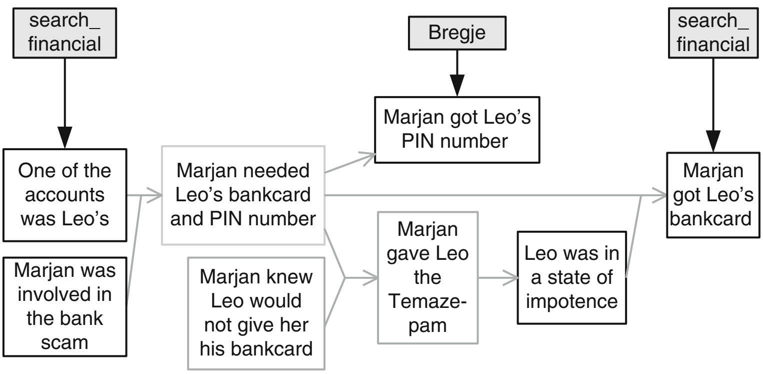



**Fig. 6.25** Marjan's financial motive for drugging Leo $S_{M\_fin}$

### 6.4.2 Marjan's Motives for Leo's Death

In the previous section, it was noted that $S_{M\_Tpam}$ is a sub episode of the larger explanation that Marjan killed Leo as Marjan wanted to bring Leo in a state of impotence is a subgoal of the higher goal Marjan wanted to kill Leo. Thus, the episode in which Marjan drugs Leo can be rendered as a sub episode just as it was in Fig. 6.23, only now the main episode is different (Fig. 6.26). Here, the exact details about Marjan giving the Temazepam and the alcohol as they appear in $S_{M\_Tpam}$ (Fig. 6.19) are summarized. The total story now fills the evidential gaps about the identity of the killer as well as the identity of the person who gave the Temazepam with hypothetical gap-fillers. However, the main episode of the story in Fig. 6.26 is incomplete (Definition 5.4.12): no motive is supplied for why Marjan wanted to kill Leo.

If we take the explanation $S_{M\_forms}$ (that Marjan wanted Leo to sign the forms for the false contract, Fig. 6.23), a possible motive for why Marjan wanted to kill Leo is that Marjan was afraid Leo might tell someone about the false contract he was made to sign once he would have sobered up (Fig. 6.27). Of course, Marjan might have thought of the fact that Leo might tell the police before she made him sign the forms. Furthermore, $S_{M\_fin}$ (that Marjan needed Leo's signature and bank card for the financial scam, Fig. 6.25) shows that Marjan maybe has an additional reason for drugging Leo. Thus, the motives for drugging Leo from $S_{M\_forms}$, $S_{M\_fin}$ and $S_{M\_kill}$ (Figs. 6.23, 6.25 and 6.26) can here be combined into one explanation that Marjan wanted to drug Leo for three main reasons: first she wanted to cloud his judgement so that he would sign the forms for the false contract; second,

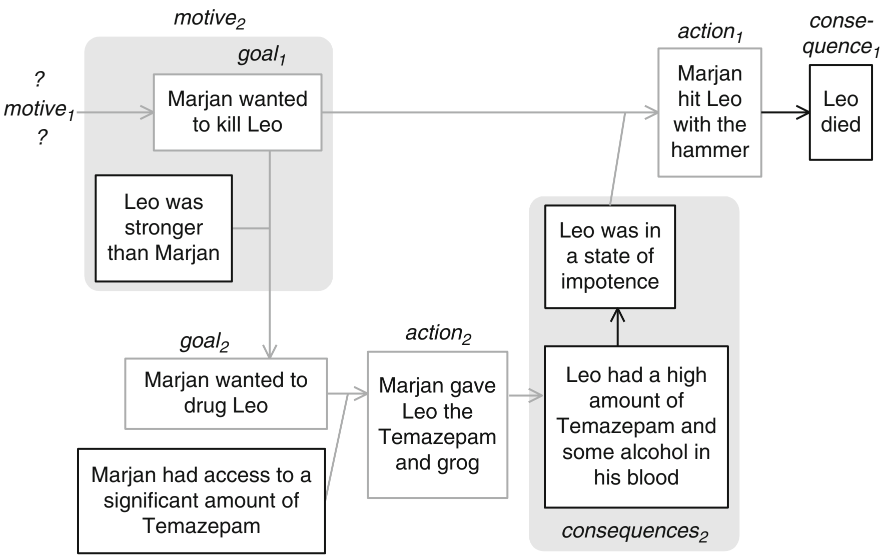



**Fig. 6.26** Marjan drugged Leo because she wanted to kill him ($S_{M\_kill}$)

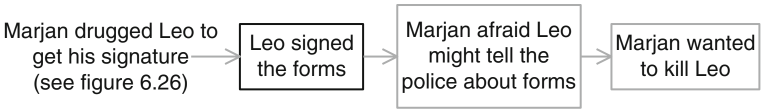



**Fig. 6.27** Marjan was afraid Leo would go to the police

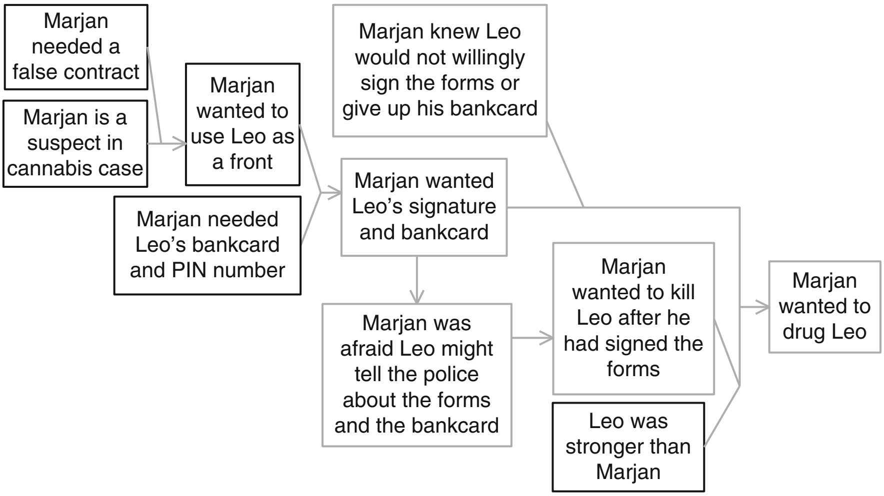



**Fig. 6.28** Overview of $S_{M\_motive}$, a combination of the motives in $S_{M\_fin}$, $S_{M\_forms}$ and $S_{M\_kill}$

she wanted to drug him so she could get his bankcard; and finally, Leo's drugged state would help Marjan kill him so that he could not go to the police. The elements of the three explanations pertaining to motives can then be combined as in Fig. 6.28.

In Fig. 6.28, not all events and causal links from the three explanations have been rendered and the focus is on the events that motivated Marjan to drug Leo. The difference between the explanations in Figs. 6.27 and 6.28 is that in the former, Marjan's goal to kill Leo is caused by Leo signing the forms and in the latter, Marjan forms the goal to kill Leo before he has actually signed the forms. The exact order in which Marjan formed her goals, however, has no big effect on the coherence or the extent to which the combined explanation conforms to the evidence (the evidential data that supports this explanation is a combination of the data supporting $S_{M\_kill}$ and $S_{M\_form}$ and $S_{M\_fin}$).

The explanation containing Marjan's motive ($S_{M\_motive}$, Fig. 6.28) can now be combined with the judiciary's story $S_{Marjan\_Jud}$ about the events surrounding Leo's death at the hands of Marjan (p. 193). The complete story containing the events up to the moment that Marjan hit Leo on the head ($S_{Marjan1}$) is then as follows:

*before the 23rd of December*

– **Marjan needed Leo's bankcard and PIN number** (employee_bank, search_fin).
– **The cannabis business is shut down by the police, Marjan is suspect** (various police reports).
– **Marjan said she rented out the barn in which the cannabis was found to a third party and that she would show the police a contract of hire** (police report).
– **Marjan asks Tasman to draft a contract** (Tasman_contract, contract).
– **Marjan wants to use Louw as a front for the business so she needs his signature** (*evidential gap*).
– **Marjan decides to drug Leo so she can get his bankcard and signature** (*evidential gap*).
– **Marjan is afraid Leo will tell others of the cannabis operation and bank fraud** (*evidential gap*).
– **Marjan decides to kill Leo when he is drugged** (*evidential gap*).

*23rd of December*

| | |
|---|---|
| before 13:00 | Leo arrived at the boarding house (MvdE_main). |
| 11:00–13:00 | Marjan (secretly) gave Leo a significant amount of Temazepam (*evidential gap*). |
| 13:00–16:30 | Waanders arrived at the boarding house (MW_main). |
| | The Temazepam had its effect on Leo, making him dizzy and numb (MW_main, tox_report). |
| | Leo signed the meaningless forms (MW_main, MvdE_main, forms). |
| 16:30–16:45 | Marjan went to town to get alcohol (MW_main, MvdE_main). |
| | Leo was in the barn for no reason, Waanders took him to the house (MW_main). |
| 16:45–17:45 | Marjan gave Leo several glasses of grog (MW_main, MvdE_main). |
| | Marjan took Leo to the hallway, "to put him in bed" (MW_main). |
| | Waanders started preparing dinner in the kitchen (MW_main). |
| 17:45–18:10 | Waanders called Marjan several times, Marjan is busy (MW_main). |
| | Waanders saw Marjan and Leo on his knees near the blue closet in the hallway (MW_main). |
| | Waanders started dinner in the living room (MW_main). |
| 18:10–18:30 | Marjan hit Leo on the head with a hammer in the hallway (location indirectly supported by DNA_1, blood_hallway; action itself is an *evidential gap*). |

Leo died because he was hit on the head with the hammer (indirectly supported by DNA_1, DNA_2, hammer, autopsy).
Leo lay bleeding on the floor of the hallway for sometime after he is hit (indirectly supported by DNA_1, blood_hallway).

This story represents a relatively well-supported and coherent explanation of Marjan drugging and killing Leo. The events that have been added to the judiciary's story from the beginning of Section 6.4 (p. 193) are in bold font.

The above story $S_{Marjan1}$ is supported by almost all the evidential data that deals with the events before Leo's death. The story is contradicted by Marjan's denial of the fact that she was involved in the cannabis case (MvdE_deny_cannabis), Marjan's denial of the fact that she gave Leo the Temazepam (MvdE_deny_Tpam) and Marjan's denial of the fact that she killed Leo (MvdE_deny_kill). This does not make a strong case against the story $S_{Marjan1}$, as there is no additional evidence that exonerates Marjan.

The main problem with the coherence of the judiciary's story on p. 208 is that it is *incomplete*. The new story $S_{Marjan1}$ adds a clear motive and goals ($S_{M\_motive}$, Fig. 6.28) and this story can be seen as completing the intentional action story scheme. However, in a story we want not only that the motive itself conforms to the evidence or our general knowledge, but also that the link between the motive and the action is supported by evidence or at least plausible given what we know about people in general. In this respect, an important causal chain in $S_{Marjan1}$ that needs to be discussed is in Fig. 6.29.

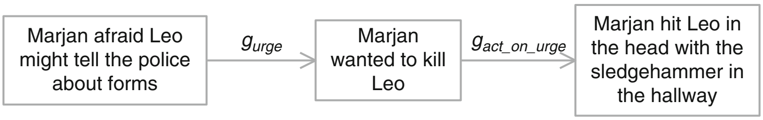



**Fig. 6.29** Marjan acting on her motive and goal

In Fig. 6.29, there are two causal relations in the chain, expressed by the generalizations $g_{urge}$ and $g_{act\_on\_urge}$. The first generalization $g_{urge}$ can be phrased as "Marjan is the kind of person who would have the urge or goal to kill Leo because she is afraid he might tell the police" and the second generalization $g_{act\_on\_urge}$ says that "Marjan is the kind of person who would act on a murderous urge or goal". Now, the story $S_{Marjan1}$ would be better if the links between Marjan's motive, goal and action as expressed in Fig. 6.29 could be supported by evidence (increasing the evidential support) or by arguments from general knowledge (increasing the (causal) plausibility).

First, consider possible arguments for the plausibility of the above generalizations. Such arguments are based on general knowledge or the reasoner's knowledge from experience. Take, for example, the arguments in Fig. 6.30. The left argument is an argument from general knowledge which supplies a reason for

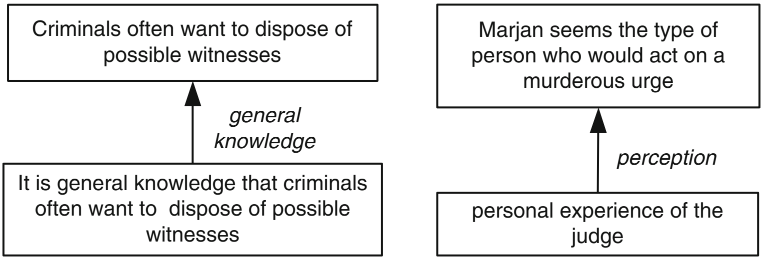



**Fig. 6.30** Arguments for the plausibility of $g_{act\_on\_urge}$ and $g_{act\_on\_urge}$

$valid(g_{act\_on\_urge})$[18] and thus improves the plausibility of $S_{Marjan1}$. The right argument in this figure is an argument from the personal experience of the judge; he or she may have observed Marjan and talked to her and found that she seems aggressive and the type of person who would be capable of killing someone. The conclusion of this argument can be seen as a reason for $valid(g_{urge})$ (which is for simplicity not rendered in the figure) and thus the argument adds to the plausibility of $S_{Marjan1}$ (Definition 5.4.7). The two arguments in Fig. 6.30 provide some basic plausibility to the story $S_{Marjan1}$, but they might not be enough to convict Marjan. If we want stronger support for Marjan's character, we would need to reason with *character evidence* about what kind of person Marjan is (Section 2.3.3). In the Court's judgment there is a paragraph detailing the findings of the psychiatrist: two of these findings are used in the arguments in Fig. 6.31. Again, the left argument can be seen as an argument for $valid(g_{urge})$ and the right argument as an argument for $valid(g_{act\_on\_urge})$. The left argument further requires us to believe an evidential

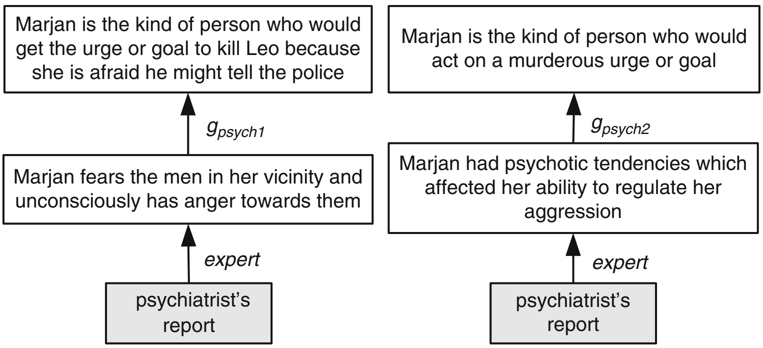



**Fig. 6.31** Supporting $g_{act\_on\_urge}$ *and* $g_{act\_on\_urge}$ with evidence for Marjan's character

[18]Formally, this argument is a reason for $valid(g_{act\_on\_urge}(x, y, c))$, where $g_{act\_on\_urge}(x, y, c)$: $x$ afraid $y$ might tell police about crime $c \Rightarrow_C x$ wanted to kill $y$. Through the reason in Definition 5.1.3 we can infer the causal generalization Marjan afraid Leo might tell police about forms $\Rightarrow_C$ Marjan wanted to kill Leo.

generalization $g_{psych1}$ ("people who have anger towards men might want to kill men if these men have incriminating knowledge about them") and the right argument requires us to believe an evidential generalization $g_{psych2}$ ("people with a low ability to regulate their aggression will act on murderous urges"). Let us assume for the moment that we accept the left argument in Fig. 6.30 for the validity of $g_{urge}$ (the general knowledge seems plausible and the generalization $g_{psych1}$ in the psychiatrist's report is somewhat far-fetched). This increases the plausibility of the story $S_{Marjan1}$. Furthermore, assume the right argument in Fig. 6.31 is accepted for the validity of $g_{act\_on\_urge}$ (the testimony by a qualified psychiatrist gives more weight to the truth of this generalization than the observations of a judge). This increases the evidential support of $S_{Marjan1}$.

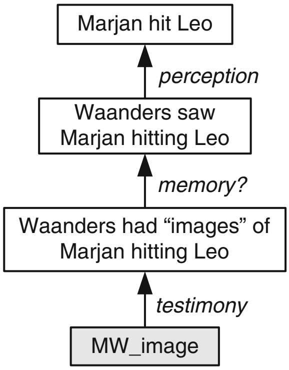
### *6.4.3 Leo's Death and the Events Afterwards*

Now that Marjan's motivation for killing Leo has been given, the events during and after his death can be discussed. These events, while they do not directly explain the explanandum (because they happened *after* Leo died), are still interesting to consider if we want the current story that Marjan drugged and killed Leo to be plausible. By providing a more elaborate story that also includes the events after Leo's death, the evidential support and coherence of the story $S_{Marjan1}$ can be improved.

In the story on p. 193, the events after Marjan hit Leo were not given, because there are multiple, sometimes contradictory, testimonies about this period. The judiciary seems to follow Beekman's and Waanders' main testimonies. However, the summarizing police report is unclear about Beekman's role in dragging the body outside and it is also not made clear exactly which of Waanders' multiple testimonies is believed. Waanders is an important witness as she was the only other person in the boarding house at the time Leo was killed by Marjan. Waanders is also a problematic witness in that she often contradicts herself, especially about crucial events. In this section, Waanders' testimonies will be considered and the course of events surrounding and just after Leo's death will be discussed. The story according to the judiciary on p. 193 will be taken as the starting point and the events and causal connectivity of this story will be discussed more or less in chronological order.

While according to the main story on p. 193, Waanders was in the living room when Leo was killed, Waanders also states that she had "images" of Marjan hitting Leo at this time. Hence we can infer that it was Marjan who hit Leo, viz. Fig. 6.32. This argument seems like a normal argument from witness testimony, which can be refined into separate inference steps for veracity, memory and perception. the question is whether the generalization that underlies the second step in the above argument, x states (s)he had "images" about p $\Rightarrow_E$ p, is based on the memory reason. For now this will be accepted. However, the statement that Waanders had "images" is vague and it can be argued that this is an indication that she misremembers (Fig. 6.33). This argument undercuts $A_{images}$. Waanders later retracts her statement and says she did not see Marjan hit Leo (Fig. 6.34). The generalization $g_{not\_see}$ is a kind of "negative version" of the perception reason: if a witness did

**Fig. 6.32** Argument $A_{images}$ about Waanders' images of Marjan hitting Leo



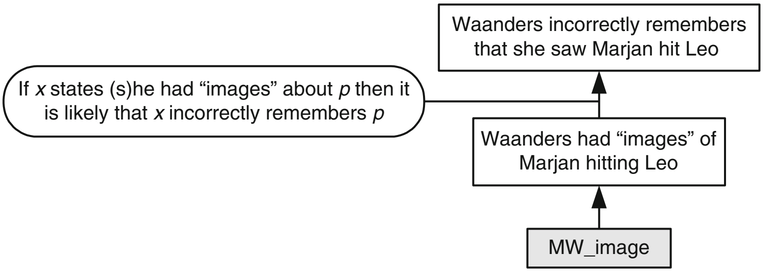
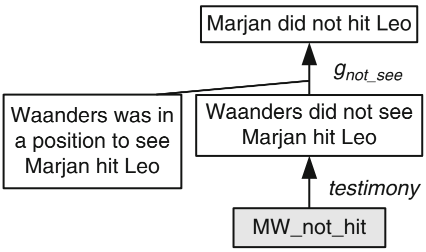




**Fig. 6.33** Waanders misremembers ($A_{misremember}$)

**Fig. 6.34** Marjan did not hit Leo ($A_{not\_hit}$ )




not see an event while the witness was in a position where she could have seen the event, then this event did not happen. This argument $A_{not\_hit}$ rebuts $A_{images}$ and vice versa.

An argument can be constructed against Waanders' veracity in either $A_{not\_hit}$ or $A_{images}$, as someone who gives conflicting testimonies may be lying, at least in one of the testimonies (Fig. 6.35). The argument $A_{conflict}$ can effectively overrule all Waanders has to say about the fact that Marjan hit Leo (i.e. $A_{not\_hit}$ or $A_{images}$). The arguments about Waanders and their attack relations can be pictured as in Fig. 6.36. Formally, the graph in Fig. 6.36 represents a part of the *evidential argumentation theory* (Definition 5.2.7), which is in turn based on Dung's (1995) ideas. Recall that

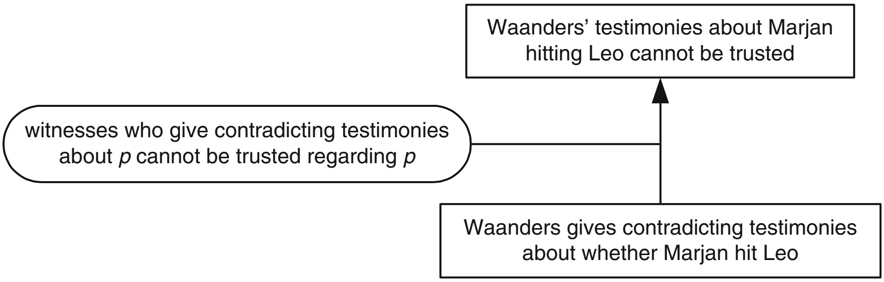



**Fig. 6.35** Witnesses are incredible if they give conflicting testimonies ($A_{conflict}$)

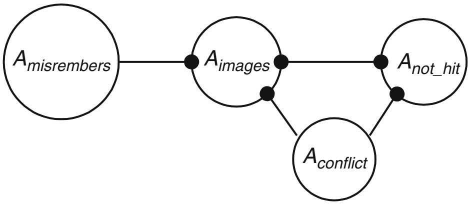



**Fig. 6.36** Waanders' testimonies about Marjan hitting Leo

given attack relations between arguments, their status can be determined (Definition 5.2.10). If $A_{conflict}$ is accepted as justified, there is nothing we can say about whether or not Marjan hit Leo when between 17:45 and 18:10. If, however, $A_{conflict}$ does not defeat the two arguments it attacks, $A_{images}$ is undercut by $A_{misremembers}$ so it is overruled and consequently $A_{not\_hit}$ is justified. For now it will be assumed that this is the case and that Marjan did not hit Leo is justifiably supported.

The current explanation $S_{Marjan1}$ assumes that Marjan hit Leo on the head some time between 18:10 and 18:30, when Waanders was eating here dinner in the living room and Beekman was still at home. Hitting someone on the head with a hammer must make quite some noise. This story-consequence can be predicted and the thus updated explanation can then be scrutinized: even though Waanders did not directly see Marjan hitting Leo, she may have heard something when Leo was killed. It actually turns out that Waanders did not hear anything while she was eating her dinner, even though it can be predicted from the explanation $S_{Marjan1}$ that she must have noticed something. Figure 6.37 shows how this story-consequence follows from two events in the main explanation $S_{Marjan1}$ and how this story-consequence can be contradicted. The explanation $S_{Marjan1}$ combined with Fig. 6.37 is now contradicted by the evidence through the generalization $g_{no\_evidence}$. This generalization, which argues that if there is no evidence for an event then the event did not take place, was already discussed before because it played a role in the argument attacking $A_{Kuilstra}$ (Fig. 6.7, p. 183).[19]

[19] Recall from the example of the dialogue game in Section 5.6 that this way of first predicting new observables and then contradicting them is not possible in the formal dialogue game.

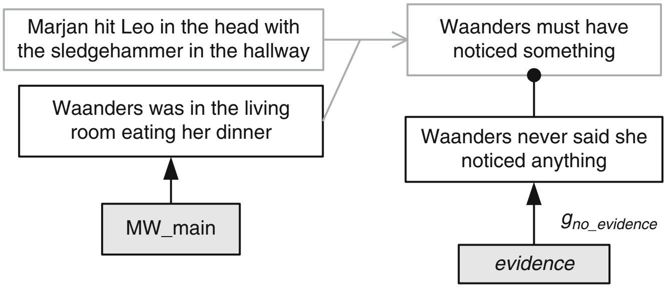



**Fig. 6.37** Predictive reasoning: Waanders must have noticed something

After Marjan allegedly killed Leo, she went to Beekman's house to tell him what she had done. According Beekman's first testimony, Marjan and Beekman knew each other because he had sold her some timber once. Beekman said that his motto was "seeing is believing" so that when Marjan came to his house and told him she killed Leo he went to check out if Marjan had really killed someone at the boarding house. This is implausible for two reasons: both Marjan and Beekman did not act in a way we expect normal people to behave (see Fig. 6.38). Notice that a large part of the story in Fig. 6.38 is inferred from Beekman's testimony. The two arguments from general knowledge are arguments against the validity of the generalizations expressing the causal relations in the story and thus add to the story's implausibility (Definition 5.4.8). Beekman later provided more information (Beekman_cannabis): he was involved in a cannabis operation with Marjan, Marjan told him she was going to use Leo as a front and she explicitly mentioned she needed his help in getting rid

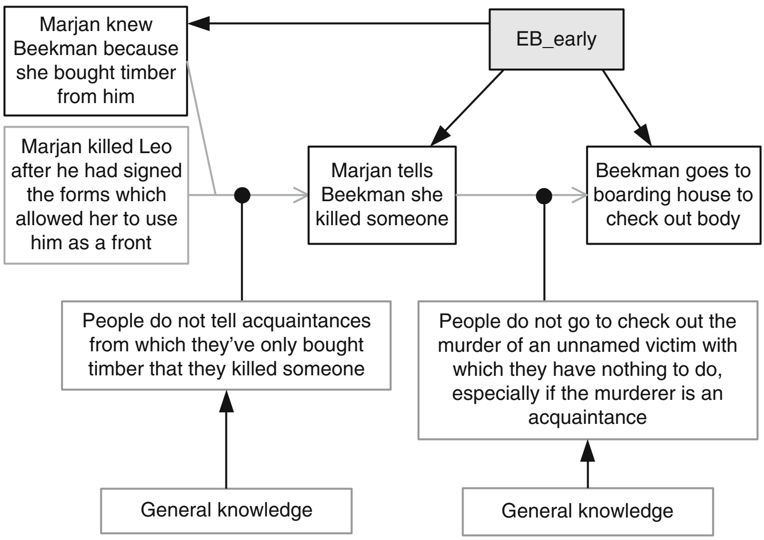



**Fig. 6.38** The implausibility of the contents of Beekman's early testimony

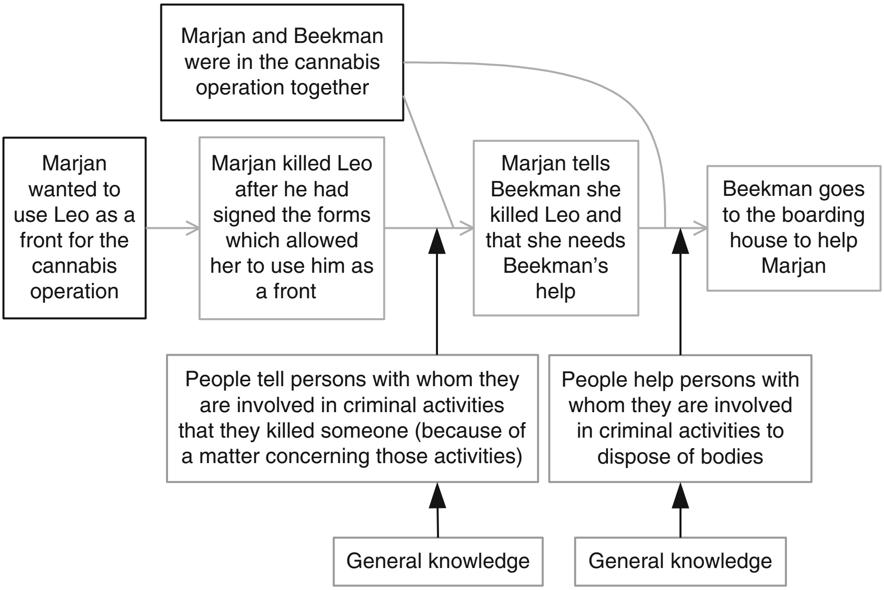



**Fig. 6.39** Beekman helped Marjan because he was also involved in the cannabis operation

of Leo when she came to his house after the murder (Fig. 6.39). This short story is more plausible than the one in Fig. 6.38: the causal generalizations are closer to what we expect than the ones in Fig. 6.38, as the explicit arguments from general knowledge show; people are much more inclined to help persons with which they run a cannabis operation than that they are inclined to help vague acquaintances. Furthermore, as Beekman knew Marjan was going to use Leo as a front, he might have guessed that Leo might have to be silenced.

There is not one testimony which definitively establishes the events from the point that Beekman arrived at the boarding house. Both Beekman and Waanders testify to slightly different accounts and Waanders also gives multiple incompatible testimonies. Waanders made an important statement about going to the toilet when Beekman was still at the boarding house. From this statement, it can be inferred that Leo was dragged outside at a point in time when Beekman was still in the boarding house (Fig. 6.40).

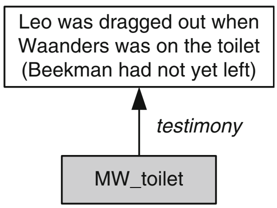



**Fig. 6.40** Argument $A_{drag}$ about the time when Leo's body was dragged outside

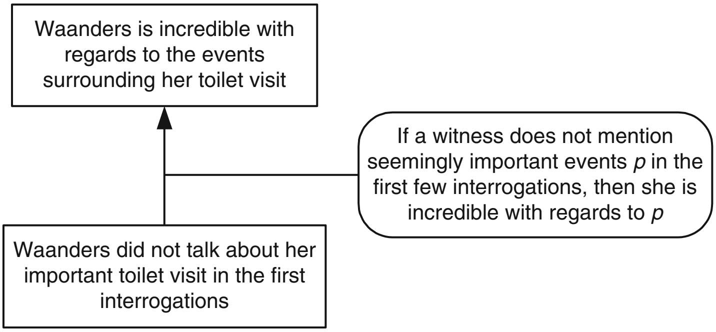



**Fig. 6.41** Waanders' incredibility about the toilet visit ($A_{incredible}$)

Waanders did not say anything about her toilet visit in the first few interrogations. This can point to an argument that she is not to be believed in her later testimony (Fig. 6.41). This argument undercuts $A_{drag}$. An argument can also be given for the credibility of Waanders statement about the toilet visit: it can be argued that the statement is credible because it is so detailed (Fig. 6.42). This is an argument about the credibility of evidence that is not based on evidential data itself. It acts as strengthening ancillary evidence to any argument based on MW_toilet_1: any attack against such an argument is attacked by $A_{credible}$. Now, the three arguments' attack relations can be rendered as in Fig. 6.43, depending on whether we accept $A_{incredible}$ or $A_{credible}$, the status of the argument for Leo was dragged out when Waanders was on the toilet can be determined. For now, $A_{credible}$ will be considered justified so that we are justified in believing what Waanders stated about her toilet visit.

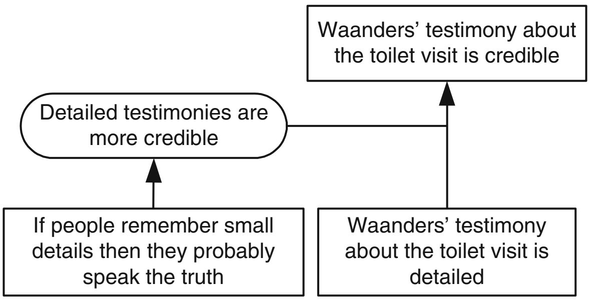



**Fig. 6.42** Waanders' credibility about the toilet visit ($A_{credible}$)

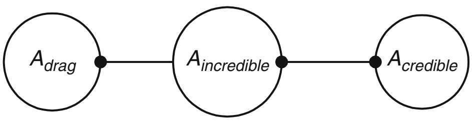



**Fig. 6.43** Waanders' testimony about her toilet visit

Now that it has been established through purely argumentative reasoning (see Figs. 6.36 and 6.43) that Leo was still alive before 18:10 (when Waanders went into the hallway to call Marjan for dinner) and that Leo was dragged out of the boarding house at the time when Waanders went to the toilet, it can be analysed whether some of the other testimonies regarding the events after Leo's death are compatible with these conclusions. An important question here is exactly when Waanders saw Leo lying in the hallway: when she went to get the washcloth (MW_hallway_1) or when she went to brush her teeth after dinner (MW_hallway_2). Accepting the second point in time leads to an inconsistent story, viz. Fig. 6.44. This short story is inconsistent because from Leo was outside it conclusively follows that ¬(Leo was in the hallway). Note that Waanders' other testimony about seeing Leo lie in the hallway, MW_hallway_1, *is* compatible with the story thus far. Hence, it can be accepted that Leo lay in the hallway (see Fig. 6.13 on p. 188) is supported by MW_hallway_1.

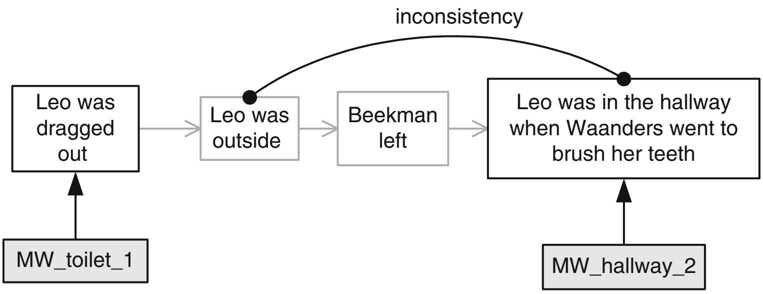



**Fig. 6.44** Inconsistent story about Leo in the hallway

The main evidential gap in the last part of the judiciary's story (p. 193) is about who exactly dragged Leo out of the front door of the boarding house (that he was dragged follows from blood_hallway and gravel_trail). In the summarizing police record, it is argued that Marjan was the one who dragged Leo out of the front door. However, this is implausible for two reasons. First, Leo is probably too heavy for Marjan to drag out on her own especially considering that, according to the judiciary's story, Marjan must have done this quickly somewhere between 18:10 and 18:30, before she went to Beekman. Second, it does not make sense that Marjan first involves Beekman in the case because she says she needs his help and subsequently drags out Leo on her own when Beekman is at the boarding house (Fig. 6.45). Notice that in Fig. 6.45, two generalizations are used to argue for the implausibility of the story, one from general knowledge and a case-specific generalization from personal experience.

One other issue that has to be discussed is Waanders' testimony that she did not see any blood. Waanders explicitly stated that she did not see any blood in the hallway, not when she went to get a washcloth, not when she went to the toilet and not when she went to brush her teeth after dinner. This contradicts a story-consequence of the story concerning the events after Leo's death (Fig. 6.46). It can be argued that

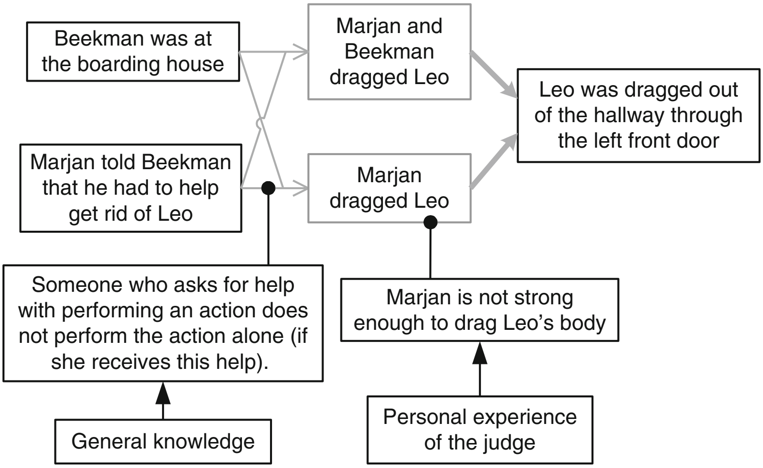



**Fig. 6.45** Beekman helped drag out Leo's body

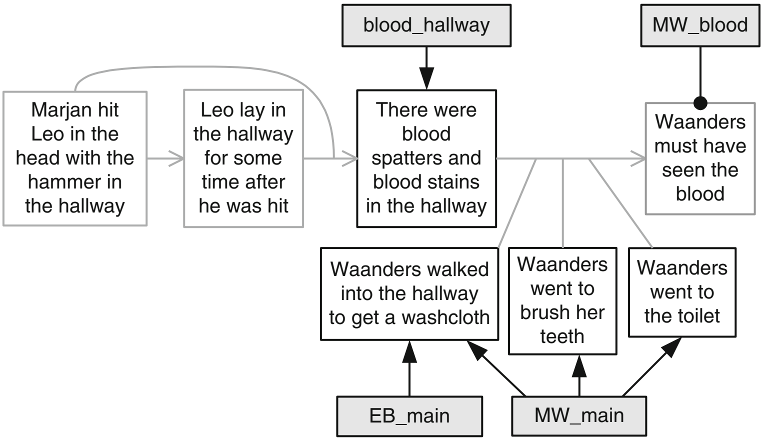



**Fig. 6.46** Waanders saw no blood

when Waanders went into the hallway for the first time to get a washcloth, Marjan had already cleaned up the worst of the blood stains. The end of the above causal chain can then be restated as in Fig. 6.47.

Beekman stated that he clearly saw traces of blood so the above story is contradicted by his testimony. Furthermore, the story may also be contradicted by the fact that the police could see some of the blood with the naked eye. However, it might also be argued that the situation was different because the police were specifically looking for traces and Waanders was not (Fig. 6.48). The main argument used the generalization $g_{easy}$: traces of blood can be seen by the police with the naked

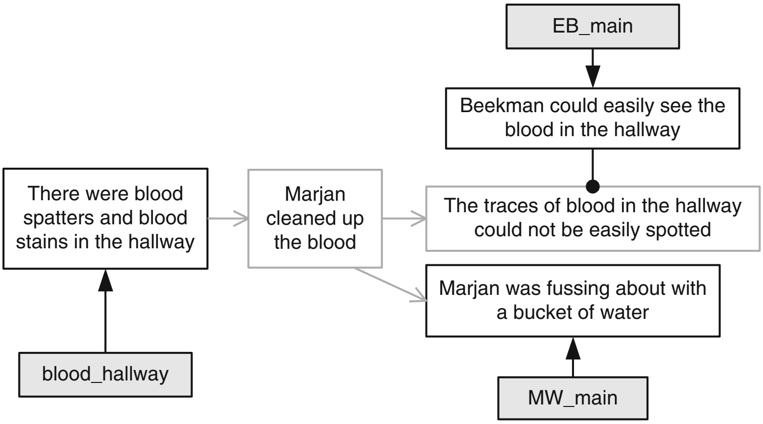



**Fig. 6.47** The blood was cleaned by Marjan

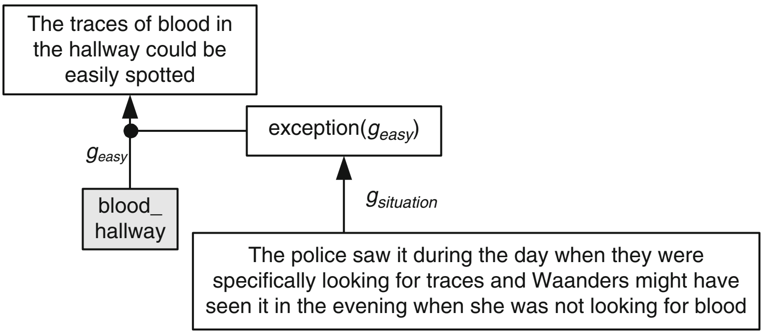



**Fig. 6.48** The traces of blood could be seen by the police

eye $\Rightarrow_E$ other possible witnesses must have also seen these traces. The undercutting argument uses the generalization $g_{situation}$: the situation in which the witness might have seen the traces and the situation in which the police saw the traces is different $\Rightarrow_E$ $exc(g_{easy})$. For the moment, Beekman's testimony will be assumed to justifiably contradict the story in Fig. 6.47 and the police's argument (Fig. 6.48) will be assumed not to contradict the story because it is overruled by the undercutter $exc(g_{easy})$.

### 6.4.4 *Summary of the Improved Judiciary's Story*

Now that the events after Leo's death have been established, the full story which has Marjan as the main perpetrator can be rendered as a combination of $S_{Marjan1}$ (end of Section 6.4.2) and the following sequence of events:

*23rd of December*

| | |
|---|---|
| 18:10–18:30 | Marjan hit Leo on the head with a hammer in the hallway (location indirectly supported by DNA_1, blood_hallway; action itself is an evidential gap). |
| | Leo died because he was hit on the head with the hammer (indirectly supported by DNA_1, DNA_2, hammer, autopsy). |
| | Leo lay bleeding on the floor of the hallway for sometime after he is hit (indirectly supported by DNA_1, blood_hallway). |
| 18:30 | Marjan arrived at Beekman's house, said to Beekman that she has "bumped off Leo" and asked him to come to the boarding house to help her get rid of Leo (EB_main, Aaltje). |
| | **Beekman, who obviously does not want any more attention to be drawn to the boarding house and cannabis operation, agrees to help Marjan** (*evidential gap*). |
| 18:45> | Beekman arrived at the boarding house and entered through the left door, he was taken to the living room by Marjan (MW_main, EB_main). |
| | Beekman and Waanders talked while Waanders was eating her dinner, Marjan left for the hallway (MW_main, EB_main). |
| | **Marjan was cleaning the blood while Waanders and Beekman talked** (*evidential gap*). |
| | Waanders went into the hallway to get a wet washcloth from the closet (MW_main, EB_main). |
| ± 19:00–19:30 | Beekman went to the hallway (MW_main, EB_main). |
| | Marjan was still scrubbing the floor of the hallway (EB_main). |
| | **Waanders went to the toilet and saw Leo being dragged** (MW_toilet_1). |
| | **Beekman and Marjan dragged Leo's body outside** (*evidential gap*). |
| | Beekman and Marjan wrapped Leo's dead body in the tent canvas outside near the front door (EB_main). |
| ± 19:20 | Beekman went home (EB_main, Aaltje). |
| 19:20–20:00 | Waanders went to brush her teeth in her room (MW_main). |
| 1:30–2:00 | Beekman went back to the boarding house (EB_main). |
| | Beekman and Marjan dragged Leo's body to the front garden (EB_main, body_Leo). |

Again, the significant differences from the judiciary's original story $S_{Marjan\_Jud}$ (p. 193) are in bold font. The combined story $S_{Marjan}$ (which is comprised of $S_{Marjan1}$ and the above events) is a coherent version of the judiciary's story $S_{Marjan\_Jud}$. It is supported by most of the evidential data in the case and is only contradicted by Marjan's allegations that she had nothing to do with all this. In Section 6.6, the quality of $S_{Marjan}$ will be further discussed when the story is briefly compared with alternative stories.

## 6.5 Alternatives to the Judiciary's Story

In the preceding Section 6.4, the judiciary's view on the case was elaborated upon and improved. This has provided a coherent story that conforms to the evidence. However, in order to safely draw a meaningful conclusion about the case, this story should be compared with alternative stories. Interesting alternative stories about the case, which are discussed in this section, concern the identity of the person who gave Leo the Temazepam and the identity of the person who killed Leo.

### *6.5.1 Leo Took the Temazepam Himself*

In the story $S_{Marjan}$, it is, for various reasons (Section 6.4.1), assumed that Marjan was the one who gave Leo the Temazepam. Except for Marga Waanders, there are no other suspects who had the physical opportunity to secretly drug Leo by giving him the Temazepam and, as was already discussed earlier, Waanders has no motives for any crime. However, a valid alternative story is that it was not Marjan who gave Leo the drugs but that Leo took the Temazepam himself, viz. Fig. 6.49.

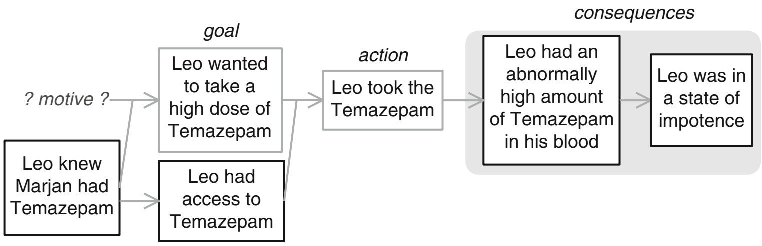



**Fig. 6.49** Leo took the Temazepam himself

From the files of the case, it follows that Leo often visited Marjan, so he might have known that Marjan had Temazepam and he might have gone and searched for it on the morning of the 23rd. Thus the states Leo knew Marjan had Temazepam and Leo had access to Marjan's Temazepam follow from the evidential data files of the case and Sturmans_Tpam, loose_Tpam, respectively. The question here is *why* Leo took the Temazepam: the explanation in Fig. 6.49 is incomplete in that it does not provide a proper motive for Leo's goal and action.

One explanation that Marjan gave in her commentary in Crombag and Israëls' book is that Leo was addicted to Temazepam. No direct evidence for this is mentioned in the book or in the judgement, but perhaps the argument in Fig. 6.50 can be constructed. The causal generalization $g_{dose}$ stands for "Someone who is addicted to Temazepam wants to take a high dose". The evidential generalization "someone who takes many types of prescription drugs is probably addicted" ($g_{drugs}$). That Leo took many drugs follows from Waanders' testimony, the fact that the Omeprazol was found and possibly also from the case file.

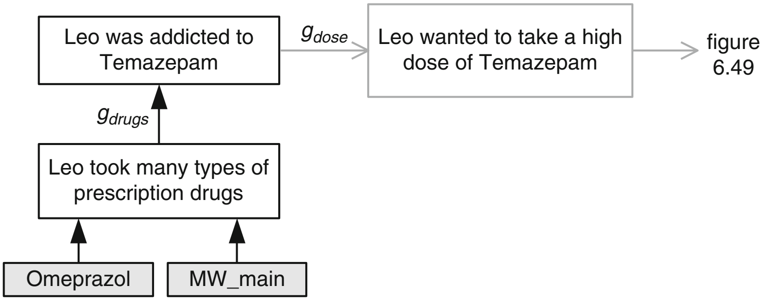



**Fig. 6.50** Leo was addicted to Temazepam ($S_{addict}$)

It is also possible that Leo tried to commit suicide. From the case file, it follows that Leo is an unstable character who has tried to end his life with sleeping pills before. Evidence points to the fact that Marjan wanted him to leave the holiday home, which he rented from her (Fig. 6.51). The argument $A_{home}$ uses a case-specific evidential generalization: "if Marjan locked the home and did not want Bregje (Leo's girlfriend) in the house, then it is likely that Marjan also wanted Leo to leave". Now, assuming that Leo had to leave the home, he is in a fairly dead-end situation: Leo is unemployed and cannot get benefits if he does not have a permanent residence address. He somehow knows that there are large quantities of Temazepam in the boarding house, so he takes it in order to try and end his life (Fig. 6.52). In Fig. 6.52, $g_{suicide}$ stands for "If Leo is in a dead-end situation, he may decide to kill himself". The validity of this generalization is supported by the fact that this is not the first time Leo tried to take his life with sleeping drugs.[20]

Considering the two explanations in which Leo took the Temazepam himself, we can probably say that $S_{addict}$ is of a relatively poor quality. First, the data Omeprazol

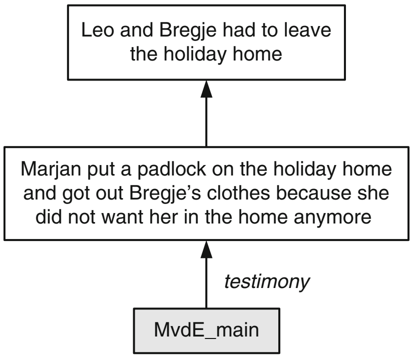



**Fig. 6.51** Marjan wanted Leo out of the holiday home ($A_{home}$)

[20]Here, the *evidential support* of the story is improved because evidence (from the case file) is given for *valid*($g_{suicide}$). The causal generalization $g_{suicide}$ is inferred through an intermediate conclusion in the argument Leo had tried to kill himself with drugs before.

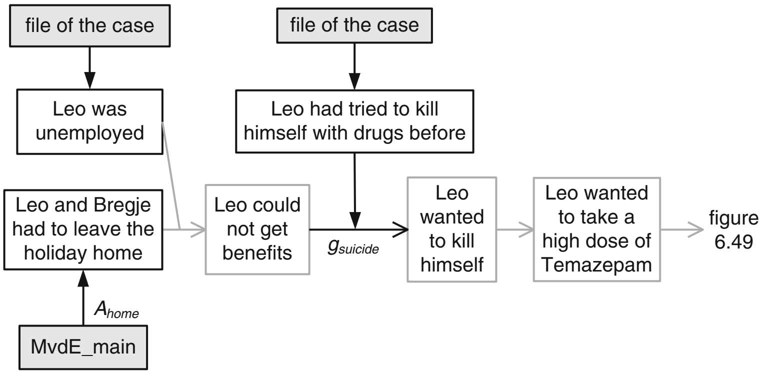



**Fig. 6.52** Leo tried to commit suicide ($S_{suicide}$)

and MW_main lends some evidential support to the explanation. However, the argument which is used to connect this evidential data to the story depends on $g_{drugs}$, a generalization for which it can be argued that it is not very strong. The inference based on this generalization can be undercut. For example, we could argue for an exception to $g_{drugs}$ (i.e. $exc(g_{drugs})$): "normally heavy users of prescription medicine get addicted, but the medicines Leo (e.g. Omeprazol) takes are not addictive" or we could argue for the invalidity of $g_{drugs}$ (i.e. $\neg valid(g_{drugs})$): "while some heavy users of prescription medicine might get addicted, this is by no means something that happens often".

We can also argue for the invalidity of the causal generalization $g_{dose}$. For example, "even addicts do not normally take the absurdly high dose Leo ingested" is a reason for $\neg valid(g_{dose})$. If this reason is based on general knowledge (i.e. "it is general knowledge that even addicts do not normally take the absurdly high dose Leo ingested"), it counts towards the implausibility of $S_{addict}$. If, for example, an expert on addictions testifies that addicts do not take such high doses, the evidential contradiction of $S_{addict}$ is increased. Whatever the source of the reason for $\neg valid(g_{dose})$, the quality of the story $S_{addict}$ can be worsened by providing this reason "even addicts not normally take the absurdly high dose Leo ingested".

The explanation $S_{suicide}$ is quite well-supported by evidential data. The pivotal issue in this story is whether the fact that Leo's unemployment, housing problem and inability to get benefits would be enough for him to try and kill himself; in other words, whether we believe $g_{suicide}$. Because there is evidence in the case that Leo had a suicidal nature,[21] the quality of the story $S_{suicide}$ is reasonable and at least as good, if not better, than the quality of $S_{addict}$.[22]

[21] This is referenced to in (Crombag and Israëls, 2008) but no further details are given so I do not know how serious Leo's previous suicide attempts were.

[22] Even if the evidential support of the two explanations is roughly the same, $S_{addict}$ is either more contradicted or less plausible because of the attack on $g_{dose}$.

An interesting question is now whether we believe that Marjan drugged Leo or that Leo took the Temazepam himself. In Section 6.4.1 it was argued that Marjan drugged Leo because she wanted his signature for the false contract ($S_{M\_forms}$, Fig. 6.23), she wanted his bankcard ($S_{M\_fin}$, Fig. 6.25) and she wanted to kill him ($S_{M\_kill}$ Fig. 6.26). The explanation $S_{M\_motive}$ (Fig. 6.28) combines these motives. Even if we decide not to believe Beekman's statement that Marjan was involved in the cannabis-growing business, there are still various independent witnesses and documents that support this (e.g. Tasman, the Mandersman affair). Furthermore, it is implausible that Marjan was completely unaware of the fact that a huge number of cannabis plants were grown in the barn next to her house.[23] Marjan's involvement in the financial scam is also supported by clear evidence. That Marjan wanted to drug Leo because this would allow her to kill him is not directly supported and depends on who we believe to have killed Leo; after all, if we say that Beekman worked alone and Marjan was unaware of Beekman's intentions to kill Leo, it makes no sense that she drugged Leo to kill him. This will be further discussed below.

As for the possibility that Leo took the Temazepam himself, recall that at the end of Section 6.5.1, it was argued that the explanation $S_{suicide}$ (that Leo wanted to commit suicide using the Temazepam) is the most likely of the alternatives involving Leo. This explanation's main strength is that it is supported by Leo's suicidal nature. The biggest problem of the suicide explanation at the moment is that it is hard to integrate it into the bigger story about Leo's death: it is fairly certain that Leo was killed by another person, as it is implausible that someone would commit suicide by first taking Temazepam and then bashing his own head in with a hammer. So the episode that Leo took the Temazepam to try and kill himself cannot be *causally connected* to the episode that Leo was killed with the hammer by either Marjan or Beekman. After all, intuitively we would regard it as an unlikely coincidence that Leo failed to kill himself with the Temazepam and that at the same moment somebody else decided to kill him because of what he knew about the cannabis operation. For the moment, it is accepted that the most likely explanation is that Marjan gave Leo the Temazepam because she wanted him to sign the forms ($S_{M\_forms}$, Fig. 6.23) and because she wanted his bankcard ($S_{M\_fin}$, Fig. 6.25). The motive $S_{M\_kill}$ (Fig. 6.26) that Marjan drugged Leo because she wanted to kill him, will be further discussed below.

[23]The implausibility of a story that Marjan did not know anything about the cannabis can be shown by, for example, using causal generalizations about people's knowledge of the contents of their barns and generalizations about the smell and energy consumption of a cannabis-growing operation to predict that Marjan must have seen something and subsequently attacking this explanation (cf. Fig. 6.37 for a similar line of reasoning).

## 6.5.2 *Beekman As the Killer*

Beneath Fig. 6.18 (p. 192) it was argued that Beekman is the only plausible other suspect for killing Leo. Interestingly, the judiciary never seems to have taken this alternative seriously. While it was considered proven that Beekman was involved in the cannabis operation and that he helped Marjan drag Leo's body, both the summarizing police report and the Court of Appeal's judgement make no mention of such an alternative scenario. Crombag and Israëls, however, propose two scenario's in which Beekman was the killer. In this section I will propose essentially three similar scenarios in which Beekman is the killer. In the first scenario, Beekman and Marjan decided that Leo should be killed together but Beekman was the one who hit Leo on the head when he came to the boarding house around 19:00. The second scenario is similar but argues that Beekman killed Leo in the night, when Marjan and Waanders were going for a walk with their dogs. The third scenario also assumes that Beekman killed Leo when the women were walking their dogs but argues that Marjan did not know Beekman was going to kill Leo.

### 6.5.2.1 Beekman Killed Leo in the Early Evening

Given the account of events, it is perfectly possible that not Marjan but Beekman was the one who killed Leo. Between 19:00 and 19:30, Marjan and Beekman were alone with Leo in the hallway and there is no direct evidence that Leo was killed before that; Waanders' testimony that it may have been Marjan is considered overruled (Fig. 6.36, p. 207). The new story $S_{Beekman1}$ can be summarized as follows:

*before the 23rd of December*

Marjan needed Leo's bankcard and PIN number.

- The cannabis business is shut down by the police, Marjan is suspect.
- Marjan said she rented out the barn in which the cannabis was found to a third party and that she would show the police a contract of hire.
- **Marjan and Beekman decide to use Leo as a front for the business.**
- Marjan decides to drug Leo so she can get his bankcard and signature.
- **Marjan and Beekman were afraid Leo will tell others of the cannabis operation and bank fraud.**
- **Beekman decides to kill Leo when he is drugged.**

*23rd of December*

| | |
|---|---|
| 11:00–18:00 | Marjan gave Leo the Temazepam and made him sign the forms. She also gave Leo several glasses of grog. Marjan then took Leo to the hallway, "to put him in bed" and left the unconscious but still alive Leo in the Hallway. |
| 18:30 | **Marjan arrived at Beekman's house to tell him that he should come and help with killing Leo.** |

| | |
|---|---|
| 18:45> | Beekman arrived at the boarding house and talked with Waanders for a while.<br>**Marjan was cleaning up Leo's vomit with a bucket.**<br>Waanders went into the hallway to get a wet washcloth from the closet. |
| ± 19:00–19:30 | Beekman went to the hallway.<br>**Beekman hit Leo on the head with a hammer in the hallway.**<br>Leo lay bleeding on the floor of the hallway for sometime after he is hit.<br>Marjan and Beekman dragged Leo outside.<br>Marjan scrubbed the floor of the hallway.<br>Beekman and Marjan wrapped Leo's dead body in the tent canvas outside near the front door. Later that night, Beekman returned and they dragged Leo's body to the lawn. |

While this story is largely the same as the $S_{Marjan}$ there are a few crucial differences, indicated in bold, which will be briefly discussed. First, in the new story Marjan does not kill Leo, so any causal chain that incorporates this event must be changed. In cases where the causal relation is not dependent on the identity of the person who hit Leo, this is trivial; for example, hitting Leo always causes blood to flow no matter who the killer is. In other cases, where the causal chain has to do with the motivations of the person who hit Leo, this new causal chain should be analysed and assessed. For example, in the new story Marjan's motives and goals no longer cause her to kill Leo and Beekman's goals and motives for killing Leo must now be made clear. Second, the evidential support and contradiction changes as the new story conforms to some evidential data that the original contradicted and vice versa. Below, first the important changes in the causal connectivity and coherence will be discussed and then the evidential support and contradiction of the new story will be assessed.

The important causal chain that was discussed in Fig. 6.29 (p. 203) expressed by Marjan was afraid Leo might tell the police about the forms and the bankcard ⇒C Marjan wanted to kill Leo $\Rightarrow_C$ Marjan hit Leo on the head with a hammer is no longer part of the explanation $S_{Beekman1}$. The first step may still be part of the explanation, as it is still possible that Marjan wanted Leo dead for the reasons specified. If we believe that Beekman killed Leo, then the second step is obviously no longer part of the explanation. Rather, a new causal link Beekman wanted to kill Leo $\Rightarrow_C$ Beekman hit Leo on the head with a hammer is now part of the story. This link says something about Beekman's character in the same way that Marjan wanted to kill Leo $\Rightarrow_C$ Marjan hit Leo on the head with a hammer says something about Marjan's character. An argument from stock of knowledge or evidence can be provided in the same way as in Figs. 6.29 and 6.30 on pp. 203, 204.

Accepting that Beekman is the kind of person who would act on his goal to kill someone still does not provide a motive for the goal that he wanted to kill Leo. Marjan's motive had to do with the bank fraud and the cannabis operation. There is no evidence that points to the fact that Beekman was also involved in the bank fraud. However, Beekman was involved in the cannabis growing operation. It might be the case that Marjan told Beekman she needed Leo done away with and that Beekman complied with Marjan's request to kill Leo because it was also in Beekman's interest that Leo would not go to the police. Or Marjan and Beekman may have decided together that Leo needed to be killed (Fig. 6.53, p. 222). Here, the causal link between Marjan and Beekman decided to kill Leo and Beekman kills Leo is important: this link can only be deemed plausible if we believe that Beekman would not let Marjan kill Leo and that Beekman agreed to kill Leo himself. This is only plausible if we agree that Beekman is a ruthless person who would do everything to protect his own freedom and that Beekman does not want to depend on Marjan to do the dirty work. Whilst the evidence in the case showed that Beekman was involved in all kinds of criminal activities, it is unclear if this evidence directly supports Beekman's ruthlessness. The assumption can also be based on personal experience (Fig. 6.54).

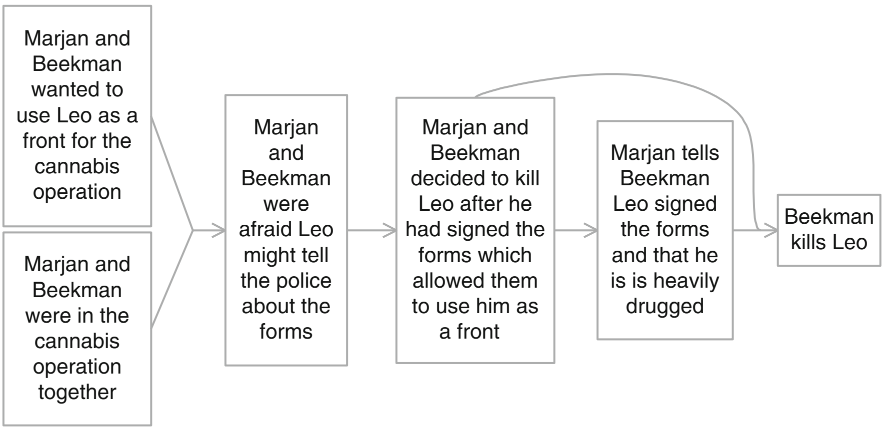



**Fig. 6.53** Beekman and Marjan's plan as Beekman's motive

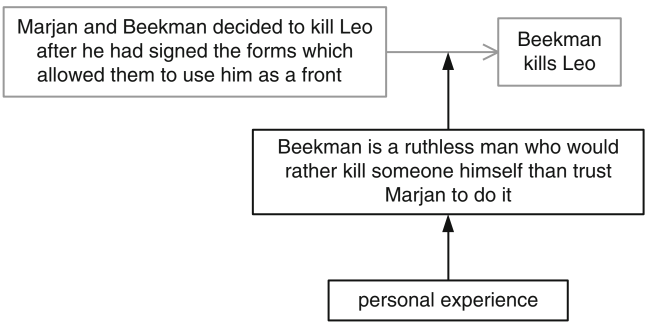



**Fig. 6.54** Beekman wanted to kill Leo himself

Accepting the new story has the consequence that Beekman's testimony EB_main now no longer totally supports the story. While there are a few events which are still supported by this testimony (for example, that Waanders went to get a washcloth or that Beekman returned in the night to drag Leo to the lawn), most of the events which are mentioned by Beekman did not happen or happened differently For example, in $S_{Beekman1}$ Marjan did not say that she killed Leo but that the details surrounding the signature had been arranged and that Leo was heavily drugged back at the boarding house. The other events to which Beekman testified between the moment Marjan came to his house and the moment that they wrapped the body in the tent canvas are also unsupported by EB_main. Waanders' testimonies MW_hallway_2 and MW_toilet_1 are compatible with the new story, because it does not matter whether Marjan or Beekman killed Leo for Waanders to have observed the things she testified to. This, however, also means that the story-consequences regarding the sound that hitting Leo must have made and the blood in the hallway are still unsupported (see Fig. 6.37, p. 208 and Fig. 6.46, p. 213). Marjan's loose statement that she did not kill Leo does not actively support $S_{Beekman1}$; however, unlike $S_{Marjan1}$, this new story is not contradicted by the statement.

Note that in the story $S_{Beekman1}$ Marjan is certainly not innocent: even if she did not agree with Beekman that Leo had to be killed, she is still involved in the cannabis operation and the bank fraud. Furthermore, she also helped Beekman to try get rid of the body.

#### 6.5.2.2 Beekman Killed Leo at Night

One of the interesting evidential contradictions for both of the stories is the fact that Waanders did not hear anything and never saw any blood. While there may be valid reasons Waanders did not see any blood (Fig. 6.47, p. 213), it does seem strange that Waanders noticed nothing if Leo was killed in the early evening when Waanders was freely walking around the boarding house. Of course, it may be that Waanders did notice something but that she has some reason for not telling this: she may to protect her friend Marjan or she may be too afraid of Beekman to say anything. Another option is that Waanders never actually heard anything or saw any blood because Leo was not killed between 18:00 and 20:00. Waanders and Marjan left the boarding house around midnight to go for a walk with their dogs. It may be that Marjan took Waanders for this walk on purpose so that Beekman could kill Leo and quickly clean the hallway. We then have a story, $S_{Beekman2}$, which is similar to $S_{Beekman1}$ only with a different time of death for Leo:

*before the 23rd of December*

Marjan and Beekman need a front for the cannabis operation and they decide to use Leo. They also decide that Leo should maybe be killed after he has signed the forms.

*23rd of December*

| | |
|---|---|
| 11:00–18:00 | Marjan gave Leo the Temazepam and made him sign the forms. She also gave Leo several glasses of grog. Marjan then took Leo |

| | |
|---|---|
| | to the hallway, "to put him in bed" and left the unconscious but still alive Leo in the Hallway. |
| 18:30 | **Marjan arrived at Beekman's house to tell him that Leo is heavily drugged and asks what do to.** |
| 18:45> | Beekman arrived at the boarding house and talked with Waanders for a while. |
| | **Marjan was cleaning up Leo's vomit with a bucket.** |
| | Waanders went into the hallway to get a wet washcloth from the closet. |
| ± 19:00–19:30 | Beekman went to the hallway. |
| | **Marjan and Beekman dragged Leo to his room.** |
| 0:00–1:00 | **Marjan and Waanders go for a long walk with the dogs.** |
| | **Beekman goes to the boarding house and finds Leo in the hallway.** |
| | **Beekman hits Leo on the head with a hammer in the hallway.** |
| | Leo lay bleeding on the floor of the hallway for sometime after he is hit. |
| | **Beekman dragged outside.** |
| | **Beekman scrubbed the floor of the hallway.** |
| | **Beekman wrapped Leo's dead body in the tent canvas and dragged Leo's body to the lawn.** |

The motives and goals of Marjan and Beekman in this story $S_{Beekman2}$ are largely the same as in $S_{Beekman1}$. A slight difference is that in $S_{Beekman2}$, it is not entirely clear why Beekman comes over to the house in the early evening. It may be that Marjan and Beekman did at first think of killing Leo. After Marjan made Leo sign she did not know what to do with him and went to Beekman. Beekman and Marjan then conferred about what to do and decided it would be best if Leo was killed later that night. Marjan deliberately took Waanders for a walk that night so that Beekman could come to the boarding house and kill Leo.

The evidential support and contradiction is similar to that of $S_{Beekman1}$: Beekman's testimony no longer supports the story and Marjan's statement no longer contradicts the story. Waanders' testimonies are still compatible with this story $S_{Beekman2}$, as it is perfectly possible that Leo was not dead when Waanders saw him lying in the hallway or when she saw him being dragged away. The main improvement of this explanation is that the fact that Waanders did not notice anything does not have to be reasoned away like in Fig. 6.47 on p. 213.

#### 6.5.2.3 Beekman Worked Alone

A totally different option is that Beekman decided to kill Leo without Marjan's consent. He must have done this when Marjan and Waanders were out for a walk at night. A large part of this new story is now the same as $S_{Beekman2}$. The main difference would be that Marjan now came to Beekman's house around 18:30 for a different reason. She may have told him that Leo was heavily drugged and that he needed to be moved or she may have told him that everything was arranged with regards to the false contract for the barn. If we believe that Marjan is not the kind of person who would want Leo dead and Beekman is, the best story is one where Beekman returns later that evening without Marjan or Waanders noticing, that he kills Leo in the hallway, drags him outside and quickly cleans the hallway.

However, the implausibility of this story can be shown using predictive reasoning. Even if Beekman killed Leo without either Marjan or Waanders knowing, the two women must have noticed the blood in the hallway when they came back from their walk or else the following day.[24] Therefore, the story that Beekman worked alone can be considered the least plausible of the alternatives in which Beekman is the killer. For the present discussion, I will assume that $S_{Beekman2}$ is the best of these alternatives.

## 6.6 Comparing the Alternatives

A comparison of alternative stories can be done informally by asking the relevant critical questions (see Section 4.4) or the stories can be compared according to the ordering given in Section 5.5.5. Recall, however, that this strict ordering is mainly to be used to guide the formal dialogue game and not to make important final decisions about which story is the best. In the preceding sections, the critical questions and their associated formal criteria were used in this way, to guide the analysis. The extent to which the explanations conform to the evidence and the coherence of the explanations was extensively discussed for each of the alternatives. As was already argued before, the ultimate decision about the quality of stories often involves substantive arguments that pertain to the content of the evidence and explanations.

Ultimately, there are two main rival stories that can be considered: either Marjan killed Leo with Beekman's help ($S_{Marjan}$, Section 6.4) or Beekman killed Leo with Marjan's help ($S_{Beekman1}$, $S_{Beekman2}$, Section 6.5.2). All these stories have some variants (e.g. regarding Marjan's precise motive) but generally speaking they are the best-supported and coherent stories about what happened on the 23rd of December 1997. The evidence in the case does not provide a decisive answer which of these possible explanations is the best: Marjan and Beekman contradict each other and the only witness who might give a decisive answer as to who killed Leo, Marga Waanders, gave incoherent and unclear statements. Therefore, the decision in the case leans heavily on the coherence of the stories $S_{Marjan}$ and $S_{Beekman2}$. This coherence in turn depends on what we believe to be "plausible" (in the literal sense as opposed to the meaning of plausible according to the formal criterion (Definition

[24]This way of predictive reasoning was illustrated in Fig. 6.37 and can be applied similarly to a case where Marjan and Waanders must have noticed something the following day.

5.4.7)). More specifically, it depends on what kinds of persons we believe Marjan and Beekman to be.

I will not expound on my own opinion about who killed Leo[25] and I will therefore not further discuss the plausibility of the two explanations $S_{Marjan}$ and $S_{Beekman2}$; interested readers are referred to Crombag and Israëls book, which is unfortunately only available in Dutch. One final interesting observation about the contents of the case is that Beekman correctly avoided one of the pitfalls of rational reasoning in the process of proof: he supplied a reasonably coherent story about what happened. Marjan's position, on the other hand, seems weaker because she does not provide a single coherent story.

## 6.7 Evaluation

The case study presented in this chapter was mainly intended to see if the hybrid theory can cope with a complex case and to show the strengths and weaknesses of the hybrid theory. The case itself is interesting because it requires both causal reasoning with stores and evidential reasoning with arguments. Having said that, the main focus in this case is on the plausibility of alternative stories and not on atomistic reasoning with evidential arguments. Since there are only a few witnesses and the main issues in the case (i.e. the identities of the person who drugged Leo and the person who killed Leo) are evidential gaps, stories play a major role in providing plausible circumstances for the assumption that either Marjan or Beekman killed Leo. The credibility and authenticity of individual evidence, which is best tested using evidential arguments, was in this case less of an issue.[26] A strong advantage of the hybrid theory when analysing and modelling a case is that it allows a flexible choice between arguments and stories as analytic tools. Often, an issue can be modelled in both a causal as well as an evidential way (see the beginning of Chapter 4 on arguments and stories as communicating vessels) and the hybrid theory allows for both. The story-based part of the hybrid theory provides an overview of the various alternatives. Imagine, for example, that only the evidential arguments in this chapter are given and that the alternatives about what happened are not written out; this would provide the puzzle pieces but no example of what the final image represented on the puzzle should look like. The stories allow us to form an idea of how the events in the case progressed and the evidential gaps point to further avenues of investigation. Both explanatory and predictive causal reasoning with (small) stories is shown to be an important in a complex case such as the Anjum case; the case involves reasoning about physical causation (see Section 6.3) as well as motivational causation (see Section 6.4 about Marjan's motives and Section 6.5.2 about

[25] One reason for this is that this case study is a simplified version which does not take into account the murder of Herre Sturmans.

[26] Interested readers are referred to Kadane and Schum's (1996) case study of the Sacco and Vanzetti case, in which issues such as witness and expert credibility are extensively discussed using (Wigmorean) argumentation.

Beekman's motives). Reasoning with motives, goals and actions is further facilitated by the story scheme for intentional actions, which shows which parts a hypothetical story are still missing. Evidential arguments are also an integral part of the case analysis: they allow us to support and attack the stories in a natural way. Evidential arguments can be used to reason about the plausibility of causal (motivational) relations in detail (see Figs. 6.29, 6.30 and 6.31 on pp. 203, 204, 205) and are thus suitable for reaching a clear consensus about the general commonsense knowledge used in the case. Furthermore, arguments allow for detailed reasoning about issues such as witness credibility (see Section 6.4.3 on Waanders' testimonies) and comparing arguments given their mutual attack relations shows how they depend on one another (see Figs. 6.36 and 6.43 on pp. 207, 211).

The procedural approach provided with the formal dialogue game is a natural way to analyse a case and the various possible speech acts and protocol rules correctly model the process of inquiry, in which alternatives are constructed, supported and contradicted. The formal criteria for the quality of a story provide clear guidelines as to how a story can be improved or worsened; modelling a particular issue causally or evidentially does not seem to affect the use of the criteria in this way. However, the ordering of explanations according to these criteria is greatly influenced by the modelling choices one makes. If, for example, every statement Beekman makes is modelled as a separate piece of evidential data, the evidential support of $S_{Marjan}$ is high, while if his main testimony is modelled as a single piece of evidence, the evidential support of $S_{Marjan}$ is low. This is not just the case for the current hybrid theory, however, but for any formal theory that allows for the comparison of positions.

The case study also shows the shortcomings of the formal dialogue game. One relevant issue here is that the game does not allow for correct predictive reasoning towards contradictory evidence (see Sections 5.6 and 5.7). As this kind of reasoning was decisive in analysing some of the main alternative stories in the case, a complete formal dialogue game should include it. Another interesting subject which deserves further research is the turntaking condition: enforcing that turns automatically change at the moment the current player's explanation is better gives the dialogue an unnatural stop-start character. As was shown in the case study, it might be a better idea to allow players to freely argue and explain and change turns at the moment they decide; this provides a more natural dialogue.

To summarize, the "static" hybrid theory is robust and flexible enough to handle large cases. Detailed discussions about evidence or single events will usually be organized and analysed using evidential arguments and status assignments, whilst story schemes and the possibility to attack and defend motivational causal relations allow for a detailed discussion about motives and character. The flexibility provided to the analyst by the combination of stories and arguments allows for the discussion of a broad range of topics in the case in the preferred; the analyst is free to choose stories or arguments are used (recall stories and arguments as communicating vessels, Chapter 4). Here the standpoint of the analyst can play an important role. For example, when not much evidence is available to the analyst, it makes sense to hypothesize one or more stories to steer the investigation. On the other hand, when the main goal is to organize the reasons for and against a single (important) witness' credibility, arguments may be more suitable. Furthermore, the role and personal taste of an analyst may also influence the analysis: some analysts (e.g. a lawyer preparing a closing statement for a jury) are perhaps more inclined to use stories whilst others (e.g. a law professor organizing the evidence in a historically important case) might use mainly evidential arguments.

As for the dynamic hybrid theory, the dialogue game can be further developed to allow for a more natural discussion about a case. However, the main underlying principles of the dialogue game, such as the way in which the formal criteria for stories steer the reasoning, provide clear guidelines for a proper rational analysis.

# Chapter 7
# Related Research on Reasoning with Criminal Evidence

This chapter concerns research related to the hybrid theory for reasoning with criminal evidence as presented in this book. Because the various preceding chapters touch on a large number of subjects,[1] a selection has been made and only the research that presents some complete model of reasoning with criminal evidence will be discussed. Research which is only relevant for some part of the current theory (for example, formal models of argumentation) is referred to in the relevant section. The chapter is divided into three parts; in Section 7.1, the (informal) theories on reasoning with stories and criminal evidence as presented by Bennett and Feldman, Pennington and Hastie and Crombag, van Koppen and Wagenaar will be discussed. Section 7.2 discusses Wigmore's charting method and Modified Wigmorean Analysis of the New Evidence Theorists. The authors discussed in these two sections will be familiar to the reader, as their theories and ideas stand at the basis of the current work. From Section 7.3 onwards, formal theories of reasoning with evidence will be discussed. Section 7.3 discusses two influential formal theories of inference to the best explanation, namely Thagard's theory of explanatory coherence (Section 7.3.1) and Josephson and Josephson's logical abductive model (Section 7.3.2). Section 7.4 discusses a prototype support system for police investigation and its underlying theory, which were developed by Keppens and a number of collaborators (Shen, Schafer, Aitken and Lee). The theory underlying the system is also a formal theory for inference to the best explanation like Thagard's and Josephson's theories. However, because of the different aims Keppens and colleagues have with their theory, the theory and the system are discussed in a separate section. Section 7.5 discusses probabilistic (Bayesian) reasoning and Bayesian Networks. This section mainly concerns Schum and Kadane's detailed work on probabilistically modelling the Sacco and Vanzetti case and also briefly discusses Hepler, Dawid and Leucari's ideas on visualizing evidence using Bayesian Networks.

[1] In particular criminal evidence (defeasible), argumentation, stories, logic and dialogue models.

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_7, © Springer Science+Business Media B.V. 2011

## 7.1 Reasoning with Stories

In the nineteen seventies, researchers in cognitive science took an interest in stories and in particular in story understanding and formalization. This research mainly focused on formal grammars for describing and analyzing stories and story understanding. One of the first to describe a general story grammar was Rumelhart (1975), which was very loosely based on Propp's analysis of folk tales (Propp, 1968). Rumelhart identified a number of basic elements of a story (e.g. Event, Setting) and defined the structure of stories by providing a set of rewrite rules. One of the general hypotheses behind this work was that information which is organized according to some grammar is easier to store in memory and recall from memory. In other words, stories which are structured according to a certain grammar should be easier to recall than stories which are not structured. This hypothesis was confirmed in experiments in which the recall of structured stories and non-structured stories was tested (e.g. Mandler and Johnson, 1977; Thorndyke, 1977). However, some researchers (Black and Wilensky, 1979) argued that story grammars were too rigid and simplistic. Often, it is not (just) the specific place an event has in a story that determines whether we can easily interpret this event (and the story as a whole), but rather knowledge of similar events or stories that influences our understanding and recall of stories. A similar idea was used by Schank and Abelson (1977), whose research concentrates on general world knowledge in the form of scripts or story schemes.

Owing in part to this work in cognitive psychology, researchers from legal psychology also became interested in stories and story-telling. The courtroom is a good example of an environment where stories play an important role, both as a tool for convincing others and a way of organizing the evidence. In Section 3.2, the research in legal psychology (Bennett and Feldman, Pennington and Hastie, Crombag, van Koppen and Wagenaar) was already discussed in the context of a general story-based approach to reasoning with evidence. Here the individual approaches will be briefly summarized and evaluated. The composition of this section is essentially chronological. First, Bennett and Feldman's and Pennington and Hastie's work is briefly discussed and then Crombag, van Koppen and Wagenaar's theory of Anchored Narratives is discussed.

### *7.1.1 Bennett and Feldman and Pennington and Hastie*

Bennett and Feldman (1981) were among the first researchers who analysed stories in the context of a criminal case. They were interested in what kind of organizational structure forms the basis of judgement in legal trials. Because of the jury system, this structure had to make all aspects of the case easily understandable for people who have no real training in dealing with legal issues. By analysing the way decision makers in criminal trials reason about a case, Bennett and Feldman found that judicial decision making depends on the construction of various stories around the available evidence. They also found that the structure of a story influences a story's believability and proposed story-based rhetorical strategies for the participants in a criminal case.

In Bennett and Feldman's model, a story is organized around a central action and the rest of the story should act as the context to this central action. Background knowledge about the world allows us to establish connections between the central action and the other elements of the story. These connections are based on our world knowledge and they are essentially commonsense generalizations (although Bennett and Feldman do not use this term). For example, "a husband and wife sleep in the same bed" or "if a car is driving then probably its engine is running". Here, the connections are not just causal relations but also empirical relations that have been established through experience or categorical relations about which concepts can be grouped together. When all the connections have been established, we end up with a network which depicts the connections in the story. In Fig. 7.1, the Haaknat story has been rendered in this way (Bennett and Feldman present similar figures, see their 1981, pp. 46, 76, 81). Here, the event "robs supermarket" is the central action and the rest of the elements are connected to this central action. The connections can be expressed by generalizations, for example "drug addicts need money".

According to Bennett and Feldman, a believable story is consistent, complete and structurally unambiguous. With completeness, Bennett and Feldman mean that a believable story must conform to some general model of social action. The five elements of scene, act, actor, agency and purpose together form a frame for the story: the *actor* was at the *scene*, where he/she used *agency* (which can be an object but also some kind of behaviour) with a *purpose* to accomplish an *act*. These elements have to be linked: the actor was at the right scene for performing the act (actor-scene-act), the actor had the purpose of performing the act (actor-purpose-act) and the actor did indeed perform the act (actor-agency-act). These links between the elements of the story (i.e. the connections in Fig. 7.1) have to be based on "unambiguous" world knowledge (Bennett and Feldman do not exactly define when a link is ambiguous).

Bennett and Feldman tested their theory about the believability of stories: 85 students were asked to assess the truth of a number of stories that were told by other students. Some of these stories were really true (that is, the events recounted had really happened) and other stories were made-up. Some of the stories (both true stories and made-up stories) were complete and unambiguous, but others were

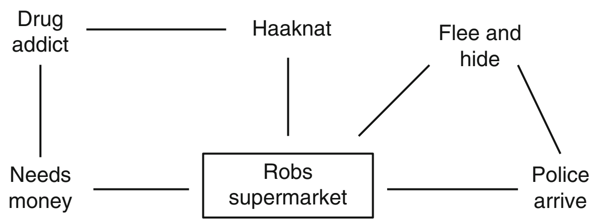



**Fig. 7.1** Graph showing the connections between different parts of the story

incomplete or had a number of ambiguous connections. It turned out that there was a significant relation between the structural completeness and ambiguity of a story and its credibility. That is, the more ambiguous connections there were in a given story, the lower its credibility was judged, irrespective of the truth of the story. So some fictional stories were judged as more credible than true stories just because they were less ambiguous.

Based on their findings, Bennett and Feldman proposed some strategies for prosecution and defence. For example, the prosecution should provide a complete and unambiguous story that the suspect committed the crime; the model of social action acts as a checklist for such a story. The defence has more options in that they only have to establish some reasonable doubt. This can be done, for example, by showing that the connections between actor-purpose-act are based on ambiguous world knowledge.

Building on the work by Bennett and Feldman, Pennington and Hastie (1986, 1988, 1992, 1993a) further developed the idea of a story in a criminal legal context. They propose a model of judicial decision making based on stories, called the Story Model of Evidence Evaluation. This model consists of three stages which represent the stages a legal decision-maker goes through when evaluating evidence: story construction and evaluation, verdict category establishment and story classification. Their model is more complex and more expressive than Bennett and Feldman's model and it incorporates additional steps, like matching a particular story to a possible verdict. This model was then extensively tested on subjects who were eligible for jury duty.

In the first stage, the evidence in the case is incorporated into a story explaining "what happened". According to Pennington and Hastie, the evidence is incorporated into one or more stories. World knowledge (e.g. knowledge about similar events) and knowledge about story structures serves to aid the juror into putting the evidence in the right causal, intentional and temporal framework. Notice that here it is the content of the evidence (i.e. the state or event $E$ and not the datum $E^*$) that is incorporated into the stories. Pennington and Hastie's view on world knowledge roughly corresponds to the current view on generalizations and the stock of knowledge (Section 2.1.3); the story structure Pennington and Hastie use is a variant on the general episode structure from Fig. 3.19 (p. 64). If more than one story is constructed, the stories are evaluated according to certainty principles that determine a subject's confidence in a story, namely coverage, consistency, plausibility, completeness and uniqueness.[2]

The second stage of the decision-making process as described by Pennington and Hastie involves learning the specific verdict categories that are applicable to the

[2]These principles were discussed in Section 3.2.4. Briefly, coverage is the extent to which the story incorporates the evidence, consistency is about the story's internal consistency and consistency with the evidence, plausibility concerns the extent to which the story conforms to the decision maker's knowledge of the world, completeness is about whether the story has all the parts of the episode scheme and uniqueness. A story is complete when all of the elements from Fig. 3.19 are part of the story.

**Table 7.1** Examples of Pennington and Hastie's verdict categories

| Verdict category | Identity | Mental state | Circumstances | Actions |
|---|---|---|---|---|
| First-degree murder | Right person | –Intent to kill<br>–Purpose formed<br>–Resolution to kill | –Insufficient provocation<br>–Interval between resolution and killing | –Unlawful killing: did not exhaust escape, excessive force<br>–Killing in pursuance of resolution |
| Second-degree murder | Right person | –Intent to inflict injury likely to result in death<br>–Deliberate, cruel act | –Insufficient provocation | –Unlawful killing: did not exhaust escape, excessive force<br>–Used deadly weapon |

decision. These categories contain the various features for each particular type of verdict. For example, the categories of first-degree and second-degree murder can be represented as in Table 7.1.

The final stage in Pennington and Hastie's model involves matching the chosen story to a verdict category. For example, if there is a story of one man killing another, the decision maker can choose between the verdict categories of first-degree murder, second-degree murder, manslaughter and self-defence. Notice how the different attributes of the verdict category correspond to the elements of the episode scheme from Fig. 3.19: the Mental State corresponds to the psychological states and goals, the Circumstances correspond to the initial states and physical states and the Actions correspond to the actions in the episode scheme. In the classification phase, the certainty principle of *goodness-of-fit* is added to the other certainty principles for choosing a story: the confidence in a story is increased the better it fits a verdict category.

Pennington and Hastie extensively tested their model. In their initial work (Pennington and Hastie, 1986), they asked subjects from a pool of volunteers eligible for jury duty to look at a re-enactment of a murder trial. Afterwards, they were interviewed and asked to say as precisely as possible how they arrived at their verdict. It was found that almost all subjects organized the evidence as a causally connected and episodically structured story of "what happened" and that information about verdicts was organized as a list of features as in Table 7.1. Interestingly, one subject tried to first identify the proposition that had to be proven and then went on to systematically look for support for this proposition, similar to formulating an ultimate probandum and supporting it with evidence through a complex argument. However, this strategy soon became too confusing and the subject finally arrived at her judgement by organizing the different elements of the case in a story.

In later work (Pennington and Hastie, 1988), the story model was further tested by requiring subjects to recognize evidence from the case. These experiments led to the conclusion that subjects spontaneously organized the evidence into stories and that the interviewing technique used in the previous experiments did not influence the results. It also turned out that the subjects better recognized the evidence that was part of the story that they had accepted as the right one. Another interesting finding was that people judged the importance of evidence according to the role the evidence played in the causal story structure: the bigger the causal role of the evidence in the story, the more important it was judged. Further experiments (Pennington and Hastie, 1992) also showed that the order in which evidence is presented influences the decision. The evidence favouring the two different positions in the case (i.e. prosecution and defence) was shown in two different orders. The evidence favouring one position was presented in a random order and the evidence favouring the other position was shown in an order which corresponded to a chronological and causal story structure. The result was that if a position was told in the chronological story order the subject more readily followed that position's verdict. Another finding was that subjects often automatically filled in certain elements of the episode structure to make complete stories. On average, a story constructed by a test subject consisted for 55% of events which were directly inferred from evidence and for 45% of events which were not inferred from evidence but added by the subjects to make a more complete story (Pennington and Hastie, 1993b).

Bennett and Feldman's and Pennington and Hastie's results show that stories play an important part in reasoning with evidence because they help people organize the evidence and make sense of a case. The experiments also show that there are dangers inherent to stories: a well-structured story is more believable than a story which is not structured, regardless of the stories' truth. Furthermore, Pennington and Hastie's finding that people better remember the evidence that was part of the story that they had accepted as the right one points to the danger of tunnel-vision, where the (provisionally) accepted story is foremost on the decision maker's or investigator's mind.

### *7.1.2 The Anchored Narratives Theory*

In 1993, three Dutch psychologists, Crombag, Van Koppen and Wagenaar, first published *Dubieuze Zaken*: *De Psychologie van Strafrechtelijk Bewijs* (*Dubious Cases*: *The Psychology of Criminal Evidence*). In 1993 an English adaptation followed (Wagenaar et al., 1993). In the book, various practices of police and judges are considered from the perspective of (legal) psychology by means of an analysis of 35 criminal cases in which the verdict was dubious. This part was subject to some (intended) controversy because the authors criticized some of the practices of police, prosecutors and lawyers in the cases they analyzed. In *Dubieuze Zaken* and later work by the authors (e.g. De Poot et al., 2004; Wagenaar and Crombag, 2005; Crombag and Israëls, 2008), a general normative theory of reasoning with criminal evidence is proposed, the Anchored Narratives Theory. It is this theory that is of interest to the current discussion.

The Anchored Narratives Theory (ANT) takes as its starting point the findings by Bennett and Feldman and Pennington and Hastie, that evidence is interpreted through a story and that the different positions in a case can be represented as stories.

A good and well-shaped story has the structure proposed by Pennington and Hastie: *initiating* or *motivating states* together with *psychological* and *physical states* lead to *goals* which lead to *actions* that cause some consequences. Furthermore, stories are hierarchically structured, which means that they can have a number of substories that elaborate on one of the elements of the story directly above it in the hierarchy. Following the experimental findings by Bennett and Feldman and Pennington and Hastie, Crombag and colleagues say that a good story is half the evidence. That is, in a case where the evidence does not provide a clear answer, a convincing story can tip the scales in favour of one of the positions in the case. However, Crombag and colleagues are well aware of the dangers of a good story pushing out a true story and hence also argue that a story should be properly *anchored* in common sense knowledge about the world.

Anchoring is best described through an example. Consider again the Haaknat case in which there were two stories explaining why Haaknat was in the moat. According to Crombag and colleagues, the prosecution's story (that Haaknat robbed the supermarket) is more believable, because Haaknat's version implies the coincidence of a man hiding in a moat exactly when and where the police are searching for a robber. This intuitive conclusion is the same as Bennett and Feldman or Pennington and Hastie would find: Haaknat's story contains more ambiguous connections (Bennett and Feldman) or is less plausible (Pennington and Hastie). In the ANT, this intuition is modelled as follows. A story is subsumed by one or more generalizations,[3] and this subsumption is called the anchoring of the story. For example, Haaknat's version of the story is only credible if we believe that, for example, "people who get into an argument usually hide from the police". According to the ANT, a story should be correctly anchored in *safe generalizations*,that is, generalizations that cannot be sensibly doubted.

Crombag and colleagues propose an analytic process through which the anchoring of a story can be improved incrementally. For example, assume that the decision maker in this case wants to analyse Haaknat's version of the story (that he had an argument with Bennie). He starts with a specific element of Haaknat's story, namely Bennie. Now, the part about Bennie in the story is subsumed under the generalization "people only lend money to people they know".[4] If Haaknat knows Bennie, he must be able to give some information about Bennie (for example, his last name or where he lives). Suppose that Haaknat gives this information which then forms a specific substory of the main story. The element "Bennie" of the main story is now anchored one level deeper in the hierarchy. At any level of a hierarchy the decision maker can decide whether or not to continue the analysis to a deeper level or to accept the current hierarchy of stories. The deeper the decision-maker descends in the hierarchy, the more specific the anchoring generalization has to be and the more

[3]Crombag and colleagues call these generalizations "common-sense rules".

[4]Note that this is not the only generalization under which this part of the story is subsumed. An example of another generalization under which this part of the story would be subsumed is "people often get into arguments about money".

specific the anchoring generalization,the more trust one can have in the story that is anchored in the generalization and the safer the decision.

Now, assume that the decision-maker decides to descend even further in the Haaknat story. The first substory of Haaknat providing Bennie's details is subsumed under the generalization "people usually do not make up details of non-existing persons". To test this generalization in the Haaknat case the police could, for example, check the address Haaknat gave. If Bennie lives at the address provided by Haaknat and Bennie confirms Haaknat's story about the fight the main story is anchored at an even deeper level. If Bennie does not live at the address or if he does not confirm Haaknat's story, there is clearly something wrong and we have what Crombag and colleagues call "unanchored evidence". In sum, Haaknat's story becomes more credible with every step taken. The anchoring process stops if the story cannot be anchored or if the decision maker decides to stop the process because the story is anchored in a safe enough generalization. Anchoring a story and its sub stories can be represented graphically (Fig. 7.2, adapted from Crombag et al., 1994, p. 72).

Each substory is a specification of the main story or one of its parts, and the substories that are deepest in the hierarchy are anchored in the "ground" of commonsense rules. As can be seen in the figure, it is possible to directly anchor a part of the main story in common knowledge; for example, the part of the story that

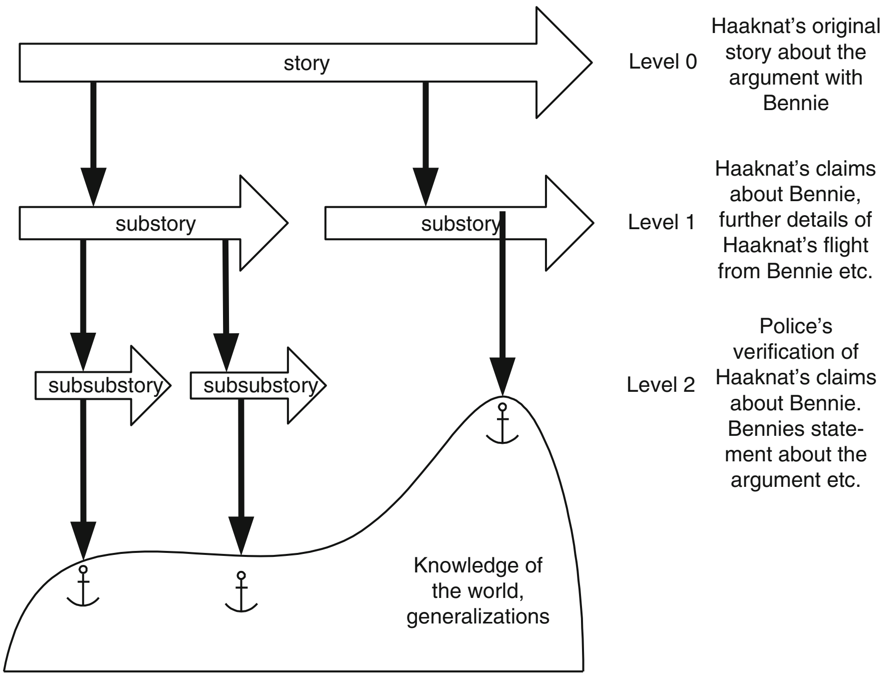



**Fig. 7.2** Anchoring Haaknat's story

Haaknat jumped into the moat can be directly anchored in the generalization "if a person jumps into a moat with water he gets wet".

Crombag and colleagues found that in many of their 35 dubious cases essential parts of the story that served as the basis for the indictment were not safely anchored or not anchored at all, and some convictions are based on unanchored stories or stories that are anchored using dubious rules or rules that are too general. The anchoring process serves as a heuristic for how people should reason with stories about the evidence and Crombag and colleagues claim that mistakes as in the dubious cases can be avoided if this heuristic is followed. However, they also recognize that in complex cases it is easier to make mistakes in the anchoring process and that therefore the heuristic may not be sufficient. Accordingly, they present their ten "universal rules of evidence" (Wagenaar et al., 1993, p. 231.), which serve to further protect the defendant against the pitfalls of the anchoring process. Note that the authors of ANT do not claim that the list of rules is exhaustive or that using these rules no more mistakes will be made; rather, the list should be regarded as complementary to the anchoring process.

1. The prosecution must present at least *one well-shaped narrative*.
2. The prosecution must present *a limited set of well-shaped narratives*.
3. *Essential components* of the narrative must be *anchored*.
4. *Anchors* for different components of the charge *should be independent* of each other.
5. The trier of fact should give *reasons for the decision* by specifying the narrative and the accompanying anchoring.
6. A fact-finder's decision as to *the level of analysis* of the evidence *should be explained through an articulation of the general beliefs used as anchors*.
7. There should be *no competing story* with equally good or better anchoring.
8. There should be *no falsifications of the* indictment's *narrative* and nested sub-narratives.
9. There should be *no anchoring onto obviously false beliefs*.
10. *The indictment and the verdict should contain the same narrative*.

Most of these rules are self-explanatory and only rule 3 and 4 will be briefly illustrated. For further qualification of these universal rules of evidence, the reader is referred to the original source.

Rule 3 states that the essential components of the story must be anchored. More specifically, at least three issues need to be explained. First the issue of *identity*, that is, is the defendant the person who committed the crime (Bennett and Feldman's *actor-scene-act*)? Second the issue of *actus reus*: what exactly did the defendant do and can this be seen as a crime (Bennett and Feldman's *actor-agency-act*; Pennington and Hastie's *actions*)? And third the issue of *mens rea*: did the defendant act intentionally or recklessly (Bennett and Feldman's *actor-purpose-act*; Pennington and Hastie's *psychological states* and *goals*)? Thus the completeness of a story is ensured. Rule 4 states that the anchors should be independent; with this, Crombag and colleagues mean that a story that serves as the basis for a conviction should not be based on one witness testimony or one confession and that ideally, a story is anchored in several pieces of evidence.

### *7.1.3 Evaluation*

Bennett and Feldman were among the first who applied general ideas about story structure and story recall to a legal setting. Their notions of completeness and connections between story elements form the basis of similar notions in other work that uses stories, including the current book. Their experimental results support the hypothesis that the structure of a story is important in judging its credibility. It should be noted that for the most part Bennett and Feldman's theory is descriptive and no normative model of the rational process of proof is given. Although some rhetorical strategies are proposed, no claim is made about the rationality of the result of their approach.

The results of Pennington and Hastie's experiments show that stories are not only a natural way of reasoning with evidence, but that there are also inherent dangers in this way of reasoning. Like Bennett and Feldman, Pennington and Hastie provide a model that is mostly descriptive: the criteria that determine the confidence in a story (i.e. evidential coverage, consistency and coherence) can be used as guidelines or heuristics in a normative theory but Pennington and Hastie's main focus is to describe how people reason with the evidence. However, their theory provides a detailed basis for any natural theory of evidence and proof.

The Anchored Narratives Theory as proposed in the work of Crombag, Van Koppen and Wagenaar is essentially the first normative theory of proof and evidence that uses stories as its basis [5] and by using ideas firmly grounded in empirical research and applying these ideas to a large number of cases, Crombag and colleagues provide a solid basis for a natural theory of evidence and proof. The idea that a story should conform to safe world knowledge (Bennett and Feldman's *unambiguous connections* and Pennington and Hastie's *plausibility*) is further elaborated upon: ANT's anchoring process provides a way of checking the extent to which a story conforms to world knowledge. The ten universal rules of evidence give a standard that should be met in a safe and correct decision about the evidence and thus also provide further heuristics that can aid in all phases of the process of proof.

In some cases, ANT is presented somewhat ambiguously and leaves room for clarification, something which was already argued in (Twining, 1995) and (Bex et al. 2006). For example, where the main story in a case is of the same form as Pennington and Hastie's general story, the substories are of an unspecified form.

[5] In this respect, ANT can be regarded as the main inspiration for the story-based approach and consequently the story-component of the hybrid theory as discussed in this thesis. In (Bex et al., 2006) the connection between ANT and a preliminary version of the hybrid theory is further discussed.

It seems to me that they can be arguments, scenarios, stories or generalizations.[6] The ambiguous presentation of ANT can cause some misunderstandings about the theory. For example, Twining (1995, p. 110) points out that the anchored narratives theory does not consider concrete evidence or case-specific generalizations. Whilst I would not make this claim, ANT does suffer from the general problem of most story-based approaches that the exact place and role of the individual pieces of evidence is not really clarified. ANT seems to consider evidence as the lowest level of substory, which is anchored in the ground of sufficiently safe commonsense generalizations. It is, however, unclear how exactly this works; in some cases (particularly their figure on p. 39 of Wagenaar et al., 1993, see Fig. 7.2) Crombag and colleagues argue that a story is anchored in commonsense knowledge whilst in other cases they say that "a story should be anchored in evidence".

Studying Pennington and Hastie's Story Model and the Anchored Narratives Theory has led me to believe that neither model is a purely story-based approach. In Pennington and Hastie's later work (particularly their 1993a), they speak of the possibility of inferring events in a story from evidence combined with world knowledge, just as in the hybrid theory. The way in which stories can be anchored in evidence using generalizations, the way of expressing exceptions to these generalizations and the dynamics of developing and refining an analysis of the evidence in a case point to an (implicit) argumentative component in the Anchored Narratives Theory.[7] As a result, both the Story Model and the Anchored Narratives Theory can be regarded as a first step towards a merged argumentative and story-based approach.

However, the hybrid theory as proposed in this book improves over both the Story Model and the Anchored Narratives Theory in several ways. Perhaps most importantly, the hybrid theory clearly specifies and disambiguates the concepts of argument and story and incorporates them both as tools in the process of proof. As was discussed earlier, stories and arguments are both necessary in the process of proof and should hence both be incorporated in a hybrid theory. For example, sometimes the credibility of a single witness or the truth of a single statement is at issue and in such cases atomistic arguments are more natural. This argumentative reasoning is only implicitly incorporated in the Story Model and the Anchored Narratives Theory. Furthermore, the evidential data in a case does not have a clear place. In the hybrid theory, the evidential data is explicitly mentioned and connected to the story through evidential arguments.

In the hybrid theory the evidence can be seen as the "ground" in which the story is anchored and the links between the evidence and the story ("anchor chains") are inferences in evidential arguments. In the hybrid theory, the way in which a story is anchored using commonsense knowledge is modelled differently from ANT. Instead of viewing generalizations or commonsense rules as the "ground" (see Fig. 7.2), a

---

[6] For example, in later work Wagenaar and Crombag (2005, p. 82) equate the generalization *after recovery of lost memories, the recollection of the recovery remains intact for a period of 9 years*' with an instance of a sub-sub-story.

[7] In recent work, Verheij and Bex (2009) have given an interpretation of the Anchored Narratives Theory that emphasizes this argumentative side of the theory.

distinction is made between internal anchors (i.e. the causal generalizations that represent the relations *in* the story) and external anchors (i.e. the evidential generalizations in the arguments from evidence, which are not part of a story). In this way, the notion of anchoring as proposed by Crombag, Van Koppen and Wagenaar is refined and it is made clearer whether we are talking about a causal story ("what happened"?) or about an individual piece of evidence ("how do we know what happened"?) when we perform the anchoring process.

The criteria for the quality of a story are more extensive and clearly defined in the hybrid theory than they are in any of the story-based models. In their rules of evidence, Crombag and colleagues provide various criteria for judging a story but do not further define them. Bennett and Feldman say nothing about evidential support or coverage and do not further define their notion of "ambiguous connections" (i.e. implausible stories). With regard to the evidence, Pennington and Hastie mention coverage (or support), contradiction and gaps but do not fully define these notions. Similarly, the notion of plausibility is not defined by them and completeness is only partly defined. Another improvement of the hybrid theory over Pennington and Hastie's Story Model is that the hybrid theory uses story schemes not only for judging the quality of a story or matching a story to a verdict category but also as general story templates in abductive reasoning.

## 7.2 Wigmore and the New Evidence Theorists

In this section I will first discuss the work by John Henry Wigmore, who can be considered as one of the "founding fathers" of the analysis of rational reasoning with evidence and proof in a legal context. After this, the work by the New Evidence Theorists Anderson, Twining, Tillers and Schum, which builds on Wigmore's work, will be discussed.

### *7.2.1 Wigmore and the Science of Judicial Proof*

Wigmore wrote some of the most complete work on evidence and proof to date. In legal circles he is best known for his major work *A Treatise on the Anglo-American System of Evidence in Trials at Common Law* (1940). This critical study of the law of evidence was also intended as a manual for practicing lawyers and law students. The Treatise, however big a success, did not contain a complete exposition of what Wigmore termed "the science of proof". According to Wigmore the study of evidence can be divided into two parts: the study of proof and the study of the legal rules of evidence. In Wigmore's day, the study of proof was ignored by most, if not all, researchers. Wigmore argued that this was a mistake, as the most important facet of the trial is the process of proof. In *The Principles of Judicial Proof* (1931), Wigmore set out to develop a broad logical, psychological and commonsense theory on the process of proof: a science of proof. His book contains extensive discussions on diverse types of evidence (e.g. handwriting, bullet marks, testimonies) and a normative, rational theory on how a mass of evidence can and should be analysed. It is this theory that is of interest in this section.

According to Wigmore, the main problem with investigation and analysing facts and the propositions that can be inferred from them is determining the effect of a mass of evidence on one's belief about a probandum, a proposition that is to be proved. An experienced lawyer or judge can quite easily draw conclusions from only a few pieces of evidence through inference and see what the possible weak points of those inferences are. However, when dealing with a mass of evidence these operations become increasingly harder. One of the reasons for this, Wigmore says, is that the human mind cannot process all the facts simultaneously: facts enter our mind over time and there are too many distractions between the moment that the fact first entered our memory and the moment that we need an old fact together with a new fact to draw a certain conclusion. In order to correctly reason with a mass of evidence, we need a way to structure and visualize this mass of evidence and the corresponding inferences and arguments. Wigmore proposes two methods, namely the Chart Method and the Narrative Method. However, his work focuses mainly on the Chart Method, as this is "the most thorough and scientific method" (Wigmore, 1931, p. 659). It involves drawing elaborate charts of nodes and links between them, where the nodes depict propositions (e.g. evidential data, probanda) and the links depict inferences between the propositions. The object of the Chart Method is for the user to consciously and rationally structure the facts and inferences in a case so that the effect of these on the ultimate probandum will become clear.

Wigmore argued for a basic logic underlying his charts and his views can be seen as a preliminary take on defeasible reasoning with arguments as discussed in Section 3.1. Wigmore distinguished between *inference* and *proof*. An inference is the process of concluding one proposition from another. An inference can be based on an explicit generalization or the generalization can be left implicit. The support that the total mass of evidence lends to the ultimate probanda is called the proof. After all the inferences offered by the proponent and opponent have been taken into account, the total effect on the probandum, the proof, can be determined. In order to combine all the inferences into a proof for some (interim) probandum, *probative processes* can be applied.

According to Wigmore there are four probative processes (Wigmore, 1931, p. 30): *assertion*, offering a premise from which a probandum can be inferred through a plausible inference; *explanation*, offering an alternative hypothesis which explains away the force of some premise for a probandum; *denial*, which stands for denying the premise from which the probandum is inferred; and *rivalry*, offering new conclusions and inferences which disprove the probandum in question. Assertion strengthens the belief in a proposition while the other processes weaken the belief in a proposition. Asserting a fact does not make it true and denying a fact does not make it false. According to Wigmore, all the evidence and chains of inferences should ultimately be weighed using experience and commonsense to see which probanda are most plausible.

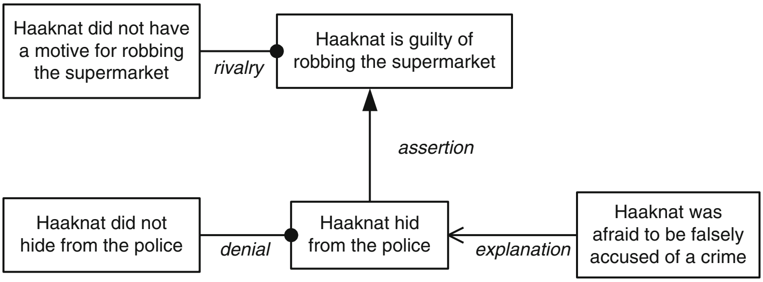
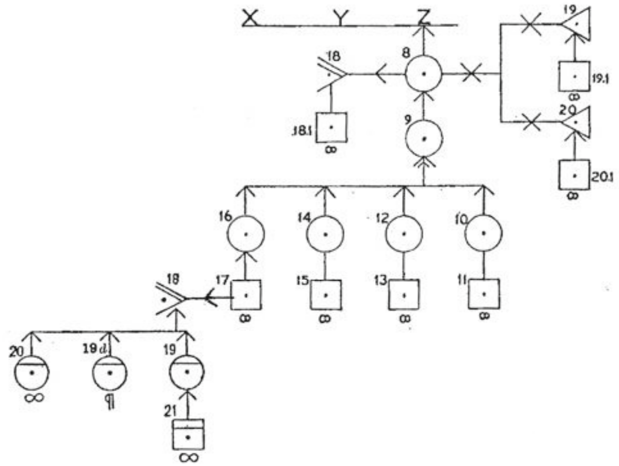



**Fig. 7.3** Wigmore's probative processes

Figure 7.3 renders the probative processes visually. An arrow with a closed head stands for "tends to prove", an arrow with an open head means "explains away" and a line with a bullet at the end means "tends to disprove". In the example, the probandum that "Haaknat is guilty of robbing the supermarket" is supported by asserting that Haaknat hid from the police. This assertion (that Haaknat hid from the police) can be denied or explained away by stating that he was afraid to be falsely accused, weakening the belief in the premise that supports the original probandum. It is further possible to attack the probandum by arguing that Haaknat never had a motive for robbing the supermarket.

As an example of an original Wigmore chart, consider Fig. 7.4 on p. 243, which has been copied from (Wigmore, 1931, p. 56). This chart concerns the Commonwealth vs. Umilian,[8] in which Umilian was accused of murdering Jedrusik. Umilian (U) and Jedrusik (J) worked on a farm together. One day, U wanted to get married and went to a priest. The priest said he could not marry U, because he had received a letter which said that U already had a wife and child. The priest investigated the letter, and it turned out that the letter was written by J and that its contents were not true. The priest then married U, but U was heard to have made threats towards J for writing the letter. A month later, J disappeared and 4 months after his disappearance, J's headless body was found in a well near the farm. U was charged with murder. In Fig. 7.4 it can be seen that the prosecution has one main argument based on four anonymous witness testimonies. The defence tries to undermine one of these by saying that the witness was biased against U. The defence also tried to explain away the revengeful emotion by saying that the emotion had faded away because eventually U did get married. The prosecution countered this by stating that U and J remained in daily contact so the wound must have rankled. Each of the nodes has been numbered and the facts corresponding to the nodes have been given in a separate list, named "key list" by Wigmore.

[8] 1901, Supreme Judicial Court of Massachusetts, 177 Mass. 582. As, obviously, no Haaknat example of an original Wigmore chart is available, this is one of the few examples not referring to Haaknat.




**KEY LIST**
**Z**: The charge that U killed J.

**8**: Revengeful murderous emotion toward J.
**9**: J's falsely charging U with bigamy, trying to prevent the marriage.
**10**: Letter received by priest stating that U already had a family in the old country.
**11**: Anonymous witnesses to 10.
**12**: J was author of letter (although it was in a fictitious name).
**13**: Anonymous witnesses to 12.
**14**: Letter communicated by priest to U.
**15**: Anonymous witnesses to 14.
**16**: Letter's statements were untrue.
**17**: Anonymous witnesses to 16.
**18**: U's marriage being finally performed, U would not have had a strong feeling of revenge.
**18.1**: Wigmore does not tell us what this represents. Maybe it is witness testimony.
**18(2)**: The witness is biased.
**19**: U and J remaining in daily contact, wound must have rankled.
**19.1**: Witness to daily contact.
**19(2)**: The witness is a discharged employee of U.
**21**: Anonymous witness to 19(2).
**19d**: Discharged employees are apt to have an emotion of hostility.
**20**: Wife remaining there, jealousy between U and J probably continued.
**20.1**: Witness to wife remaining.
**20(2)**: The witness' strong demeanour of bias while on the stand.

**Fig. 7.4** Chart for the Umilian case

The square nodes stand for testimonial evidence and the circle nodes represent interim probanda or generalizations. When a proposition is offered as evidence for another proposition (*assertion*), this is denoted by a vertical line between the two propositions; a vertical arrow means that one proposition tends to prove another with normal force; a vertical arrow with a double arrowhead means that the force is strong. Triangular nodes denote evidence that either weakens (or "tends to disprove", denoted with < through the line) or strengthens (or "tends to prove", denoted with X through the line) the proposition to which they are linked with a horizontal line. Finally, ¶ below a node means that the fact has been judicially noticed as a matter of general knowledge, and $\infty$ below a node means that the evidence was directly offered to the court. All bottom nodes of a graph should have either ¶ or $\infty$ below them, as otherwise there are facts which have not been admitted or witnessed in court and there is no evidence for them.[9] Wigmore uses a number of other symbols (e.g. different arrows for different strengths of inferences); most of them are variations on the symbols used in the example.

In addition to the Chart Method, Wigmore briefly discussed the Narrative Method of analysis (Wigmore, 1931, pp. 659–663). This Narrative Method is simpler and closer to the way evidence is usually described by lawyers. According to Wigmore, a narrative should start with an introductory statement detailing the initial clues. Then the evidence is classified in various categories, where, for example, first the evidence pertaining to Motive is recited, then the evidence relevant for Opportunity is recited and so on. After the classification of the evidence, the conclusions drawn from the evidence should be stated.

### 7.2.2 Modified Wigmorean Analysis

After Wigmore, the interest in a theory on proof and rational reasoning with evidence decreased. Most legal theorists continued their research which focused more on the law of evidence than on factual inference. However, in the last 30 years, a movement that has become known as the "New Evidence Scholarship" has become an important force in research on evidence. New Evidence Scholarship includes research and researchers who have the same interests, namely factual evidence and reasoning with this evidence in a legal context. Subjects are, among other things, the logic of inferences about facts and how to use formal probabilistic methods in evaluating evidence.[10] A recent book exemplary for the work of the New Evidence Theorists is the *Analysis of Evidence* (Anderson et al., 2005). In this book, Anderson, Twining and Schum present their Modified Wigmorean Analysis (MWA) and its connection with other methods for the analysis of a mass of evidence, such as outlines and stories. In addition to these various analytical methods, other subjects like types of evidence, probabilistic and Bayesian methods for evaluating evidence and the connection between the "science of proof" and the law of evidence are discussed. In keeping with the idea of this related-research chapter, below only the various tools and ideas used in MWA will be discussed.

Anderson, Twining and Schum argue that there are two important reasons for studying argument charts of the type that Wigmore proposed. First, they create the awareness that arguments about evidence are essentially combinations of propositions and their relations, and that in everyday discourse large parts of arguments are often overlooked. Second, the chart method forces the user to make all these different propositions and their relations and combinations explicit. This was exactly Wigmore's main reason for the method; according to Wigmore, making all the different steps of an argument explicit allows one to see the gaps and possible contradictions in an argument, and it also allows one to find weak and strong points

[9] ¶ and $\infty$ are analogous to propositions from the sets $G_E$ and $I_E$ (Section 5.2), respectively.

[10] The important work by the New Evidence Theorists such as Anderson, Tillers, Twining and Schum has already been mentioned in preceding chapters of this thesis.

in the different arguments used. According to Anderson and colleagues, the charting method is best suited to rigorous analysis of facts and arguments. More than Wigmore, Anderson and colleagues focus on the uses and dangers of generalizations, which are extensively discussed in their work (e.g. Anderson, 1999; Twining, 1999; Anderson et al., 2005, Chapter 10). MWA also proposes other methods of analysis that use outlines, chronologies and stories. It is argued that these methods serve their own purpose but that they are in general not precise enough to incorporate all the evidence and arguments in a well-organized way. Below first the chart method of MWA will be discussed and then the other methods of analysis proposed by Anderson and colleagues will be briefly discussed.

The authors of the *Analysis of Evidence* agree that the charting method can be overwhelming at first. The new user has to learn and apply a significant number of symbols and techniques which he is not acquainted with. To ease the use of charts, Anderson and colleagues proposes a protocol, which together with a more simplified charting method should facilitate the construction and readability of charts. The first step of this protocol is to clarify the standpoint of the person doing the analysis. Who am I? At what stage in the process am I? What am I trying to do? This is important because once the standpoint is clear, the person doing the analysis knows on what areas he should concentrate. If, for example, he is a lawyer preparing for trial, looking for new evidence might still be an option. A student analyzing a trial to get acquainted with the method of Modified Wigmorean Analysis cannot search for new evidence and should concentrate on other things.

The second step is to formulate the ultimate probandum or probanda. These depend on a rule or rules of law, and are often given by the indictment. The ultimate probandum is the condition of the rule of law that applies in the case at hand. Often it is a conjunction of different propositions. These are the penultimate probanda, and the object of the third step is to formulate these material facts which have to be proven separately.

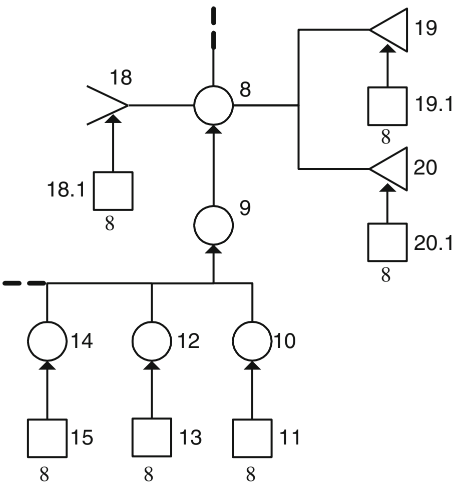
The fourth step is to formulate the provisional theories of the case. A theory is a strategic argument about the case as a whole. The theory usually has as its conclusion propositions which concern the ultimate outcome of the case, usually penultimate probanda. Two simple examples of different theories are:

> Haaknat is an addict who was in need of money and therefore decided to rob the supermarket. The fact that he was found hiding in the moat shows that he was trying to evade the police. Therefore, it was Haaknat who robbed the supermarket.
>
> During his argument with Bennie, Haaknat felt threatened by Bennie and therefore wanted to get as far from Bennie as possible. He hid in the moat because he was afraid of Bennie and it was a coincidence that the police found him there while searching for the robber.

By formulating a theory, the analyst expresses his first ideas about the case and a theory provides a starting point from which to proceed. When working on a non-decided case a lawyer should formulate a number of theories so that he has exhausted all possibilities and is not surprised at the actual trial. Multiple theories can be formulated early on in the trial and some theories will be discarded as soon as new evidence is available that contradicts the theory.

Step five involves formulating the key list, that is all the propositions the analyst expects the chart to contain. Together with step six, constructing the chart, this can be considered the most time-consuming process. The analyst works in two directions, bottom-up and top-down. Bottom-up means working from the available evidence to the penultimate probanda. Bottom-down means working from the penultimate probanda expressed in the different theories and finding out which evidence, if any, support the probanda. Finally, step seven is the evaluation of the chart. Is the chart readable? Are there any weaknesses in the arguments? This step is of course continuous throughout the whole process, but some defects can only be seen when the whole chart is finished.

Figure 7.5 depicts a part of the Umilian chart (see Fig. 7.4) which I have rendered in the simplified format as proposed by Anderson and colleagues; not much has been changed but the chart has been cleaned up and is easier to read. Here, squares denote testimonial evidence, circles denote other evidence or propositions inferred from other propositions. A closed triangle stands for strengthening ancillary evidence and an open triangle stands for weakening ancillary evidence. Note that the various probative processes have been abstracted from and that here only the strengthening or weakening effect of a proposition is considered.




**KEY LIST**
**Z**: The charge that U killed J.

**8**: Revengeful murderous emotion toward J.
**9**: J's falsely charging U with bigamy, trying to prevent the marriage.
**10**: Letter received by priest stating that U already had a family in the old country.
**11**: Anonymous witnesses to 10.
**12**: J was author of letter (although it was in a fictitious name).
**13**: Anonymous witnesses to 12.
**14**: Letter communicated by priest to U.
**15**: Anonymous witnesses to 14.
**18**: U's marriage being finally performed, U would not have had a strong feeling of revenge.
**18.1**: Wigmore does not tell us what this represents. Maybe it is witness testimony.
**19**: U and J remaining in daily contact, wound must have rankled.
**19.1**: Witness to daily contact.
**20**: Wife remaining there, jealousy between U and J probably continued.
**20.1**: Witness to wife remaining.

**Fig. 7.5** Part of the Umilian chart

In addition to the analysis of evidence using the seven step protocol and modified Wigmore charts as shown above, Anderson and colleagues also provide the outline method of analysis, which is closer to how practicing lawyers actually manage and analyse a case. The outline method shows many similarities with Wigmore's Narrative Method as described above. In this method, the probanda are listed and possible evidence for a probandum is noted as a sub list, for example:

(1) Haaknat robbed the supermarket.

   (a) Motive: Haaknat has a motive for the robbery.

      (i) Haaknat is a drug addict who was in need of money.

   (b) Opportunity: Haaknat has the opportunity to rob the supermarket.

      (i) Haaknat was last seen by a friend the day before the robbery in the same city as the supermarket.
      (ii) Someone looking like Haaknat was seen near the supermarket shortly after the robbery.

Such an outline is further supported by two analytical devices, namely chronologies and stories (or narratives). Chronologies put a certain type of events in chronological order; for example, a witness-by-witness chronology puts the events to which witnesses can testify into chronological order and a master chronology puts all events for which there is evidence into chronological order. Chronologies can be useful for identifying gaps and provide an overview of the case and they can form the basis for stories.

According to Anderson and colleagues, stories are more detailed than chronologies and form a meaningful whole because of their causal connectedness. Anderson and colleagues mention Bennett and Feldman and Pennington and Hastie when they say that stories are psychologically necessary: a story can be used to organize the evidence and stories can also be used in the investigative phase, where they serve as possible hypotheses. Tillers (2005) argues that stories (or scenarios as he calls them) can be used to insert hypothetical gap-fillers in between events for which there is evidence. Stories, however, are also dangerous vehicles for "irrational means of persuasion" (Anderson et al., 2005, p. 281). They can be used to sneak in irrelevant facts, appeal to hidden stereotypes, make use of dubious analogies and so on. Anderson and colleagues propose a protocol for assessing a story, which consists of a number of critical questions that can be asked when a story is provided in a criminal case. These questions allow for the testing of the evidentiary support, plausibility and coherence of a story. Examples of questions are "Is the story internally consistent?", "Is there evidence that conflicts with the story?", "Does the story fit some familiar story?" and "Insofar as the story goes beyond the data, is the story supported by background generalizations?".

### 7.2.3 Evaluation

Wigmore's work on proof has had a big influence on the scholars on evidence and reasoning with evidence today. The scope of his work is enormous and in this section only a small part of his ideas is summarized, in particular his ideas on charting evidence. Wigmore was a pioneer in many of the subjects discussed in this book: non-legal reasoning about evidence, reasoning with complex arguments, visualization and sense-making were all extensively discussed by Wigmore. Following Wigmore's pioneering work, the New Evidence Theorists also present a broad and detailed view on criminal evidence. Their continued research on rational reasoning with criminal evidence, of which (Anderson et al., 2005; Schum, 1994; Twining, 2006) are some of the main results, has had a profound impact on this book.

The work by Wigmore and the New Evidence Theorists is not only intended to be a theoretical exposition of the science or principles of proof, but also as a method that can be used by practicing lawyers and law students. However, Anderson, Twining and Schum make an important distinction here, namely between the use of the charting method as a teaching tool[11] and its use as a practical tool by trial lawyers. As a teaching and learning tool, the theoretical background must be sound, while the usability of the method is not the most important, because in an academic setting the mental training with the theory is most important. In practical use, other concerns such as usability and naturalness play a larger role. Twining (2007) remarks that charts have only a limited functionality in presentation and communication of arguments about evidence and that the chart method and the outline method with its chronologies and stories should be seen as complementary rather than exclusive.[12]

With the chart method and their seven-step protocol, Anderson and colleagues clearly have a normative goal just as the authors of the Anchored Narratives Theory have. Anderson and colleagues seem to follow Wigmore in their claim that the chart method is the most rational method and that the main focus should always be on the reasoning from evidential data to probanda. Stories are dangerous and the ultimate decision in a criminal case should ideally never be based on a story; stories are psychological tools to be used in organization and communication and never in rational decision making. That said, in the new edition of the *Analysis of Evidence*, Anderson and colleagues recognize that stories are necessary for organizing the evidence and in the investigative phase. It is for this reason that the above-mentioned protocol for testing stories, which is similar to

---

[11] All the authors of the *Analysis of Evidence* have used the chart method as a teaching tool in courses on their respective universities.

[12] Schum (2005) has similarly argued for stories as a way of organizing the probanda and Tillers (2005) explicitly shows the connection between the evidence and a story.

some of the criteria as proposed by Bennett and Feldman, Pennington and Hastie and Crombag and colleagues (Section 7.1), is added to the Modified Wigmorean Analysis.

Like the Story Model and the Anchored Narratives Theory, the latest version of Modified Wigmorean Analysis also takes steps towards a hybrid theory. However, arguments and stories are not fully integrated in the theory presented in the *Analysis of Evidence*. Arguments (or charts) and stories are not directly linked and in many ways, stories and arguments are regarded as different ways of presenting the same reasoning: stories play mainly a psychological role while arguments play a rational role. In this book I have argued that in the process of proof, which is inference to the best explanation, stories can be used as complex abductive hypotheses. Once the evidence has a clear place and stories can be criticized and compared using arguments, inference to the best explanation with stories becomes not only a natural but also a rationally acceptable way of reasoning.

Stories are necessary in any theory of reasoning with criminal evidence. For example, at certain times in an investigation we can use a story to fill an evidential gap and see whether assuming the gap-filling event still allows for a plausible story. Furthermore, our world knowledge is not just a collection of generalizations but also contains stories and, more importantly, story schemes which effectively "summarize" a complex chain of causal or temporal relations. Anderson and colleagues recognize that stories may be useful for purposes of investigation and a detailed outline is similar to what in this book is called a story. Tillers (2005) also explicitly discusses the gap-filling function of stories or scenarios, which he illustrates with figures similar to the ones presented in this book. The hybrid theory extends Anderson and colleagues' outline method and Tillers' scenario-based method. Arguments can be used to directly support or contradict the story (i.e. the outline or scenario) and a story's quality are depends on evidential arguments. This direct interaction between arguments and stories is unique to the hybrid theory and allows for a natural and rationally well-founded theory of inference to the best explanation.

Following earlier work by Bex et al. (2003), the current work improves on Modified Wigmorean Analysis by explicitly connecting arguments to stories and by providing a detailed formal analysis of reasoning with different types of evidence. This has yielded standard argumentation schemes or prima facie reasons for particular types of evidence, which can easily be used in the dialectical process of argumentation. The formal analysis has also provided a systematic account of (conditional and non-conditional) generalizations and various ways to attack these generalizations. In addition to aiding in the detailed and systematic analysis of reasoning with evidence, the formalization of arguments also provides us with a logical underpinning for the argument-based component of the hybrid theory. For example, the difference between reasoning about the validity of generalizations and exceptions to generalizations has been logically modelled and the argumentation-based semantics that govern the status of arguments provide a solid formal basis to the comparison of arguments.

## 7.3 Theoretical Models of Inference to the Best Explanation

This section briefly discusses two model-based theories of inference to the best explanation. First, Thagard's theory of Explanatory Coherence is discussed and then Josephson and Josephson's abductive logical theory as applied to a simple criminal case is discussed. In Section 7.3.3, the two theories are evaluated and compared with the hybrid theory from this book.

### *7.3.1 Thagard's Explanatory Coherence*

Thagard original aims for developing his theory of Explanatory Coherence (Thagard, 1989) were to provide a computational theory with which the coherence of (scientific) hypotheses can be tested and which gives a psychologically natural account of inference to the best explanation. Because the process of proof is essentially also comprised of inference to the best explanation, Thagard has applied his model to various cases concerning reasoning with criminal evidence.

Explanatory coherence takes as its basis a connectionist network in which nodes are connected by non-directed links. Each node has a (numerical) activation value which can spread through the network: neighbouring nodes connected to an activated node by an excitatory link rise in activation and neighbouring nodes connected through an inhibitory link decrease in activation. In Thagard's model, the nodes represent evidential data and hypothesized propositions and the links represent the coherence between the propositions represented by the nodes. Acceptance or rejection of a proposition is represented by the degree of activation of the node, where an evidence node is always connected to a special node with the highest possible activation. In Thagard's ECHO program, the activation of all nodes in a network is spread until a state has been reached in which the activation of the nodes no longer changes.

The connectionist algorithms used by Thagard are relatively standard. The most important part of Thagard's theory is therefore comprised of his seven coherence principles, principles that govern when two propositions cohere. Below the term *hypothesis* stands for a single hypothesized proposition (instead of a combination of propositions) and *evidence* stands for a single piece of evidence:

Principle E1 – Symmetry:

Explanatory coherence is a symmetric relation, unlike, say, conditional probability. That is, two propositions p and q cohere with each other equally.

Principle E2 – Explanation:

(a) A hypothesis coheres with what it explains, which can either be evidence or another hypothesis;

(b) Hypotheses that together explain some other proposition cohere with each other; and

(c) The more hypotheses it takes to explain something, the lower the degree of coherence between these hypotheses.

Principle E3 – Analogy:
Similar hypotheses that explain similar pieces of evidence cohere.

Principle E4 – Data priority:
Evidence, which describes the results of observations, has a degree of acceptability on its own.

Principle E5 – Contradiction:
Contradictory propositions are incoherent with each other.

Principle E6 – Competition:
If p and q both explain a proposition, and if p and q are not explanatorily connected, then p and q are incoherent with each other (p and q are explanatorily connected if one explains the other or if together they explain something).

Principle E7 – Acceptance:
The acceptability of a proposition in a network of propositions depends on its coherence with them.

Notice that the coherence relation is closely related to the *explains* relation and thus the *causal* relation (see Section 2.3.2) and the principles underlying explanatory coherence are very similar to the general principles underlying abductive inference to the best explanation and causal model-based reasoning. For example, if in a causal theory $a$ causes $b$, then according to Thagard $a$ explains $b$ and thus, according to principle $E2$, $a$ and $b$ cohere. Consequently, the symmetrical coherence relation is essentially a combination of the causal and evidential relation between two propositions. Another example is the principle of competition ($E6$), which is similar to the condition in the definition of an explanation (Definition 5.3.4) that two different causes for one event should be considered as different explanations and hence should be compared with each other. Thagard claims that a connectionist network based on the above principles can be seen as a (causally connected) *story* of which the coherence can be determined by first applying the above principles to a set of hypotheses and evidence and subsequently determining the activation values of the nodes.

In his work (1989, 2004, 2005), Thagard provides some studies of real cases modelled as a coherence network. In (Thagard, 2004), he models the well-known von Bülow case, in which Claus von Bülow was charged with killing his (wealthy) socialite wife, Sunny. The program ECHO gives as its output the same hypotheses as in the original case. That is, the networks of accepted hypotheses (or the accepted stories) are the same as the juries accepted in the original case.

## *7.3.2 Josephson's Logical Model of Abduction*

In MacCrimmon and Tillers' collection of short papers (MacCrimmon and Tillers, 2002), John Josephson applies his general views on logical abductive reasoning to a legal case in which reasoning with evidence is performed. Josephson's approach[13] has much in common with Thagard's approach: a hypotheses, which stands for a single proposition, may *explain* some data or observations. Hypotheses may also imply or explain each other and can be combined to explain data. They can also be incompatible (Thagard's *contradiction*) or they may be independent (Thagard's *competition*), which means that they are not necessarily incompatible but provide alternative explanation for some data. Finally, each hypothesis has a confidence score which is one of nine values ranging from *ruled-out* through *neutral* to *confirmed*.

In the article, Josephson gives an example of an update strategy that determines the best explanation in a knowledge-based system for abductive reasoning. This system is in many ways similar to Thagard's ECHO[14]: a case is represented as a network of hypotheses and evidence with positive links (*explains* or *implies*) and negative links (*incompatible*). Each of the hypotheses and evidence nodes has an initial confidence score and the scores are propagated through the network by a repeated cycle of updates, as follows:

1. Find all essential hypotheses, hypotheses without which the important evidence cannot be explained.
2. Hypotheses that are strongly incompatible with these essential hypotheses should be deleted and hypotheses that are weakly compatible should have their confidence score lowered.
3. New essential hypotheses may come up as their competitors are deleted; with the new essential hypotheses, start the process again at 1.
4. Repeat until no new essential hypotheses are discovered.

Ideally, all the data is now explained. If this is not the case, the system starts a similar cycle in which the essential hypotheses are not those that explain all the data but rather those that explain as much of the data as possible.

Josephson gives an example of an actual legal case containing 17 pieces of evidence and 16 possible hypotheses. The results provided by the system represent the best explanations given the way in which the case is modelled. Josephson argues that the evidence is weighed in a reasonable way and that the process detailed above takes advantage of various ways in which hypotheses can interact.

### *7.3.3 Evaluation*

Both Thagard's and Josephson's approaches discussed in the above section are closely related to the abstract story-based approach as detailed in Section 3.2 and

[13] As developed in (Josephson and Josephson, 1994). See Sections 2.3.1 (abductive reasoning), 3.2.3 (explaining the evidence), and 4.4 (comparing stories) for Josephson's views on abductive reasoning.

[14] Josephson's system is not based on connectionist network algorithms but rather on Josephson's own "Essentials First, Leveraging Incompatibility (EFLI) strategy" (cf. Josephson and Josephson, 1994).

both models are based on the same principles of causal-abductive inference to the best explanation that also underlie the causal part of the hybrid theory.

Thagard claims that his theory of explanatory coherence is psychologically plausible (here he refers to Pennington and Hastie's work), that it is computationally plausible and that it is closely related to several ideas from the philosophy of science such as holism (Quine and Ulian, 1970). In this way, Thagard presents a theory that meets the requirements set in Chapter 1 by meeting cognitive, computational and conceptual standards. Through a large number of case studies not only in a legal context but on scientific theories or common sense hypotheses, Thagard claims to have shown that the connectionist model correctly describes how humans perform inference to the best explanation. Thagard also claims that his principles of coherence not only describe human reasoning, but that they can also serve as rules for correct rational inference to the best explanation. However, he does not further expound on this remark in the context of the process of proof and reasoning with evidence.

Whilst the initial activation of the nodes has to be set as a numerical value and the weights for the excitatory and inhibitory links have to be input by the user, Thagard claims that the actual values do not matter much as long as the values for the excitatory weights are higher than those for the inhibitory weights. Thus his approach does not suffer from the "number problem", where complex probabilities or activation values have to be guessed in order for the model to be meaningful and to give proper results. However, a significant disadvantage of Thagard's model is that the complex numerical algorithm employed makes the process of determining the coherence of an explanation somewhat of a black box: it is not always clear why one set of coherent hypotheses is more strongly activated (and thus more coherent) than another.

Josephson's experimental work focuses on determining what kinds of reasoning can and should be used when determining the best explanation for some evidence. His model as applied to the case study provides interesting insights about inference to the best explanation and also provides a few possible heuristic rules for performing such inferences (for example, start with the hypotheses that are essential and refine from there). Josephson's model computes the best explanation by using discrete confidence levels as opposed to the real numbers in Thagard's connectionist model. This allows for more insight into the actual process of selecting the best hypotheses. However, it is unclear whether in Josephson's approach the confidence levels influence the final result; if this is the case, guessing these confidence levels is an important component of modelling a case and the modeller is faced with the above-mentioned "number problem".

Because Thagard's and Josephson's theories essentially model the purely story-based approach, both models have the advantages and disadvantages of such an approach. For example, advantages are that Thagard's an Josephson's theories are close to how humans actually reason with evidence and that they allow for a relatively simple overview of the evidence and the hypotheses in a case. Disadvantages are that the position of the evidential data and its relation to the events (i.e. the hypotheses) is not always clear and that atomistic reasoning about a single piece of evidence and its conclusions is not possible. Additionally, it is impossible to reason *about* the explanatory model and the outcome of the model (i.e. the best explanation). The hybrid theory takes a procedural approach to rationality and allows causal links and links between the evidence and the explanation to be challenged in a dialectical setting. In Thagard's and Josephson's theories, this is not possible and hence the explanatory models of stories cannot be checked.

## 7.4 Keppens and Colleagues' Decision Support System for Police Investigation

Keppens and Schafer (2006) have developed a logical model-based reasoning technique specifically aimed at a decision support system for crime investigation. The motivation for their research is essentially the same as the motivation for the current work: in crime investigations the risk of tunnel-vision is real, as demonstrated by the examples of miscarriages of justice Keppens and Schafer give. The proposed system (of which a prototype has been implemented) has as its objective to aid investigators by, given some evidence, automatically constructing possible scenarios that explain this evidence and, given these hypothetical scenarios, suggesting additional evidence that could be found.

The system is based on basic principles of model-based diagnosis: given a set of causal rules $T$ of the form "*cause* → *effect*" and some observations $O$, hypothesized literals $H$ are abduced which together with the rules imply the observations: $H \cup T \vdash O$. In (Keppens and Schafer, 2006) the hypothetical scenarios consisting of the literals together with the causal rules are part of a *scenario space*. Because in crime investigation there are often a large number of possible scenarios given some evidence, the knowledge base does not contain complete scenarios but rather scenario fragments from which the final scenario space in a case is constructed. For example, the causal rule hanging(victim) → suffer(victim, asphyxiation), which denotes that someone who is hanging by his neck with a rope will choke occurs not only in the *suicide* scenario but also in, for example the *accident* scenario (the victim planned to cut the rope but failed to do so) or the *murder* scenario (someone killed the victim by hanging or made it want to appear as if the victim hanged himself). In this way, it is not necessary to have complete scenarios in the knowledge base but rather a set of causal rules which can be combined into multiple possible scenarios. Note that these scenarios concern not only what happened during the crime (or accident) but also any possible actions by the investigators; for example, a scenario may contain the causal rule suffer(victim, asphyxiation) ∧ correct_diagnosis(pathologist, cause_of_death(victim)) → medical_report(pathologist, cause_of_death(victim), asphyctiation).

After the initial scenarios have been abduced, the system predicts through deductive reasoning what other possible states could be caused by the current hypothesized scenarios. In this way, new observables can be predicted and it becomes clear what evidence might still be collected to prove or disprove these new observables. The predicted observables may be used to differentiate between the various hypothesized scenarios. In later work (Shen et al., 2007), one additional explanatory stage is added after this prediction stage, in which new causes for this uncollected evidence are abduced so that all the possible causes of the new evidence are considered by the system. In the final stage of automated reasoning, constraints are added to the scenario space, which determine which scenario fragments may occur together. These constraints are deduced from the knowledge base. For example, the rule $\mathsf{suicide} \wedge \mathsf{accident} \rightarrow \perp$ ensures that the propositions suicide and accident cannot occur in the same explanation so that the system "maintains the truth".

With each new piece of evidence, the system updates its scenario space. In between updates, the user can query the system and the system will send a report to the user. Questions that can be asked are "Which hypotheses are supported by the available evidence?", "What additional pieces of evidence can be found if a certain scenario would be true?" and "what pieces or sets of additional evidence would differentiate between the scenarios?". Answers to the last two questions effectively propose investigative actions; for example, a psychologist's testimony that the victim was actively suicidal would help to differentiate between *suicide* and *murder* hypotheses so a possible investigative action might be to find out if the victim was seeing a psychologist.

### *7.4.1 Evaluation*

Keppens and colleagues clearly discuss their motivation for and the aims they have with their model, namely that it serves as the basis for an automated decision support system which helps in police investigations. The model that underlies the system is based on the same principles of causal-abductive inference to the best explanation that also underlie the story-based approach and the causal part of the hybrid theory. Keppens and colleagues, however, make no claim as to what is the best explanation and only use their causal scenarios for determining possible next steps in the investigation. That is, the model does not return the best scenario from a set of possible scenarios but rather the evidence which might discriminate between the hypothetical scenarios. The way in which the system proposes possible avenues of investigation is useful for practical investigative purposes and such a system can be a powerful tool if used properly.

However, a knowledge-based system is only as good as its knowledge base and for Keppens and colleagues' approach to be of any real use, a complex knowledge base containing a large amount of scenario fragments must be constructed. This can be a time-consuming process which suffers from the so-called "knowledge acquisition bottleneck": experts (i.e. investigators or police analysts) may provide incomplete or even incorrect knowledge and they might not always be able to correctly articulate their knowledge. A second problem with an automated expert system might be that once the knowledge base becomes sufficiently complex, the users of the system will not be able to grasp the information contained in the knowledge base in its totality. Therefore, the user must simply trust the output of the model without having a complete view of all the evidence and all scenarios.

Keppens and colleagues seem to initially avoid these general problems by keeping the scenarios and scenario fragments small and understandable. Furthermore, the query function of the prototype system seems to return scenarios as simple outlines or lists of events and does not burden the user with the explicit causal information. However, this is also dangerous as faulty causal rules can influence the output of the system and thus the investigation. In Keppens and colleagues' approach, the causal model bases its rationality on the fact that it is constructed by experts, whereas the current hybrid approach takes a procedural approach to rationality in which the model can be challenged in a dialectical setting.

## 7.5 Probabilistic Reasoning and Bayesian Belief Networks

The use of probabilities and statistics in the courtroom is a hot topic nowadays. Forensic scientists often speak about DNA evidence using numerical probabilities (e.g. the blood evidence in the Anjum case, Table 3 on p. 168 and Fig. 6.9 on p. 185) and experts in statistics have been called upon to testify about the probability that some crime was committed given the circumstances of the case. In addition to this, sometimes controversial, use of statistics, probabilities have also served another role, namely as the basis for a formal framework for describing and quantifying the relations between evidence and hypotheses. As Dawid (2005) puts it: "the theory of probability is nothing less than the logic of inference under uncertainty". It is this second use of probability theory that is of interest to this book and in this section, some formal theories of reasoning with evidence that use (Bayesian) probability theory will be briefly discussed. First, a very brief introduction to some of the concepts in Bayesian probability theory will be given. In Section 7.5.1, Kadane and Schum's use of Bayesian probability theory to establish the force of evidence in Wigmore graphs will be discussed. In Section 7.5.2, Bayesian Networks and the use of sensitivity analysis will be briefly discussed. Section 7.5.2 discusses Hepler et al.'s use of Bayesian Belief Networks for visualizing evidence and in Section 7.5.3 the Bayesian approaches are evaluated.

The theories discussed in this section are all subjectivist theories of probabilities. Such a theory allows one to make a subjective probability estimate about the occurrence of an event without appealing to some repeated instance of a similar event. This estimate is expressed as $P(A)$ (the probability of a proposition $A$) and is between 0 and 1, where $P(\mathsf{true}) = 1$ and $P(\mathsf{false}) = 0$. Bayes' well-known theorem plays a large role in subjectivist accounts of probability, evidence and belief revision. Published posthumously (Bayes, 1763), the theorem is important in performing calculations with *conditional probabilities*. A conditional probability concerning evidence, written as $P(H \mid E)$, is the probability of the occurrence of a hypothesis $H$ given certain evidence $E$. Usually, it is exactly this probability we are interested in. That is, we want to determine what the effect of a piece of evidence on a hypothesis is and which hypothesis is the most likely given the evidence.

Bayes' Theorem is the following equation:

- $P(H|E) = P(H) \times P(E|H)/P(E)$.

It allows for the derivation of the *posterior probability* $P(H \mid E)$ given $P(H)$, the *prior probability* of the hypothesis, and $P(E \mid H)$ and $P(E \mid \neg H)$, the *likelihood* of $E$ given $H$ or $\neg H$. The term $P(E)$ can be determined by the following formula:

- $P(E) = P(E|H) \times P(H) + P(E|\neg H) \times P(\neg H)$.

Say, for example, that we want to determine $P(\text{H\_car} \mid \text{w}_1)$, the probability that Haaknat got into the car given that the witness $w_1$ saw someone looking like Haaknat get into his car. By Bayes' Theorem, $P(\text{H\_car} \mid \text{w}_1) = P(\text{H\_car}) \times P(\text{w}_1 \mid \text{H\_car})/P(\text{w}_1)$. Now, the following probabilities need to be determined: $P(\text{w}_1 \mid \text{H\_car})$, the probability that $w_1$ saw someone looking like Haaknat get into the car if Haaknat got into the car, $P(\text{w}_1 \mid \neg(\text{H\_car}))$, the probability that $w_1$ saw someone looking like Haaknat get into his car if Haaknat did not get into the car, and $P(\text{H\_car})$, the prior probability that Haaknat got into his car at that particular place at that particular moment. Say that $P(\text{w}_1 \mid \text{H\_car}) = 1$ (if Haaknat actually got into his car, the witness in all probability saw Haaknat), $P(\text{w}_1 \mid \neg(\text{H\_car})) = 0.4$ (there is a reasonable chance that either the witness saw someone else who merely looked like Haaknat), $P(\text{H\_car}) = 0.3$ (the fact that Haaknat lived close by and that the car was his makes it somewhat likely that Haaknat got into the car) and $P(\neg(\text{Hcar})) = 1 - 0.3 = 0.7$, then $P(\text{w}_1) = 1 \times 0.3 + 0.4 \times 0.7 = 0.58$. Consequently, the posterior probability $P(\text{H\_car} \mid \text{w}_1) = 0.3 \times 1/0.58 \approx 0.517$. So the witness evidence has a reasonable effect on the probability that it was actually Haaknat who got in the car: before the evidence, the probability that it was Haaknat is 0.3 and with the evidence taken into account it is 0.517.

### *7.5.1 Kadane and Schum's Analysis of the Sacco and Vanzetti Case*

In their Probabilistic Analysis of the Sacco and Vanzetti Case (Kadane and Schum, 1996), Kadane and Schum combine Modified Wigmorean Analysis with ideas on Bayesian probabilities and likelihood. They give a thorough analysis of the well-known Sacco and Vanzetti case, providing Wigmore charts detailing the evidence and reasoning in the case. At the same time, they express ideas on the probative force of individual and combined evidence in the case using Bayesian methods.

Kadane and Schum are mainly concerned with the probative force of evidence in moving from a prior belief about the probability of some hypothesis to a posterior belief about the probability of the hypothesis and they are not so much interested in the prior probabilities of events. Thus they steer clear of a major point of controversy among statisticians, namely if and how the prior probability $P(H)$ can be determined *before* any evidence in the case has been considered. For example, how can the prior probability $P(\text{H robs SM})$, that Haaknat robbed the supermarket, correctly be determined? Is Haaknat as likely as 800 other people in the neighbourhood to have committed the crime, $P(\text{H robs SM}) = 0.00125$? Should we take his background as a drug addict into consideration? If we do, can we still say we are talking about the prior probability, as this background can also be seen as evidence? These are all hard questions to answer but in their analysis, Kadane and Schum focus on the conditional probabilities and not on the priors.

Kadane and Schum particularly focus on what they call the *odds-likelihood* ratio. The odds of a hypothesis $H$ to $\neg H$ (i.e. the ratio $H/\neg H$) are denoted as $(H{:}\neg H)$ and the posterior odds of $H$ to $\neg H$ given evidence $E$ are denoted as $(H{:}\ \neg H \mid E)$. Now, the odds-likelihood ratio $L_e$ is given as $(H{:}\ \neg H \mid E)/(H{:}\neg H)$. This ratio indicates the change of odds as the result of taking the evidence into account. In the cases that $L_e > 1$, the evidence has a positive effect on the hypothesis $H$, that is, the hypothesis $H$ is favoured over $\neg H$ given the evidence. In cases were $L_e < 1$, the evidence has a positive effect on the hypothesis $\neg H$, that is, the evidence favours $\neg H$ over $H$.

In the case study of the Sacco and Vanzetti case, much time was taken up by the construction of the Wigmore charts. Kaiser (a historian with much knowledge of the case), Kadane and Schum then each provided likelihood ratio assessments for collections of evidence that corresponded to some particular Wigmore chart. In this way, they wanted to analyse what three persons knowledgeable about the case thought was the actual effect of evidence on two final hypotheses, namely "Sacco was guilty but not Vanzetti" and "Sacco and Vanzetti were innocent". In other words, the assessments involved the effect of the evidence on the odds that Sacco was guilty or not. For example, Chart 4 is about whether witnesses saw Sacco at the scene of the crime.[15] This chart contains statements by prosecution witnesses Pelser, Wade and Frantello claiming to have seen or heard Sacco just before or during the crime and statements by defence witnesses who claim that the prosecution's witnesses' veracity, objectivity or observational sensitivity is questionable, thus providing ancillary evidence that weakens the inference from the prosecution witnesses to the state that "Sacco was at the crime scene". According to Schum, the evidence in Chart 4 favours Sacco's innocence at a ratio of 2:1. The idea behind this ratio is that the defence witnesses weakening the prosecution witnesses' statements were more credible than the prosecution witnesses who directly supported Sacco's presence and that thus the prosecution's evidence had no positive effect on the hypothesis that Sacco was guilty. The positive effect on Sacco's innocence came from the argument that 34 other witnesses failed to identify Sacco. Kadane and Kaiser, however, thought otherwise: Kaiser argued that the evidence only favours Sacco's innocence at a ratio of 1:1.1 and Kadane thought neither Sacco's innocence nor his guilt were favoured by the evidence in the chart.

Kadane and Schum's analysis of the Sacco and Vanzetti case involves subjective probability guesses regarding the probative force of evidence by the authors and Kaiser. The assessments are based on each assessor's background knowledge (of the case) and the authors do not claim that their assessment is the definitive one. The assessors had two major points of disagreement, each of which involved assessing the probative force of evidence spread over 4 complex Wigmore charts concerning the complex firearm evidence in the Sacco and Vanzetti case. This shows that in complex cases, correctly judging the aggregate force of evidence is hard; the "number problem" discussed in Section 7.3.3 (correctly judging probabilities and having

[15]For reasons of space, the chart will not be reproduced here, see (Kadane and Schum, 1996, p. 290) or (Bex et al., 2003, Fig. 4).

some measure of a "correct" probability) becomes more serious the more complex the evidence and the combinations of evidence become. In the next section, a way of analysing the force of evidence which does not involve subjective probability assessments is proposed.

### *7.5.2 Bayesian Belief Networks and Sensitivity Analysis*

A useful tool in reasoning with Bayesian probabilities are Bayesian Networks or Belief Networks. Such Bayesian Networks[16] provide a graphical representation of the conditional dependencies and independencies between propositions and it can be used to compute complex (conditional) probabilities. A Bayesian Network can be used to visualize evidence and inference and provides a tool with which complex probabilistic calculations can be performed. Examples of the use of Bayesian Networks in reasoning about evidence are (Kadane and Schum, 1996; Levitt and Laskey, 2000; Thagard, 2004). In this section, Bayesian Networks and their main uses will be briefly discussed.

A Bayesian Network consists of two parts, a qualitative part and a quantitative part. The qualitative part is a directed acyclic graph, of which every node denotes a proposition (e.g. a hypothesis or a piece of evidence). This graph is a graphical representation of the conditional dependencies and independencies holding among the nodes. Consider, for example, the graph in Fig. 7.6.

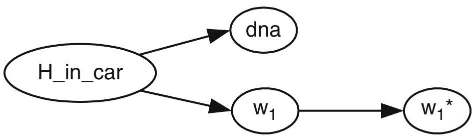



**Fig. 7.6** A simple Bayesian network

This graph denotes the dependencies between $w_1$* (evidence, the statement that witness $w_1$ saw someone who looked like Haaknat), $w_1$ (the event that $w_1$ saw someone who looked like Haaknat), dna (the fact that the DNA of the hair in the car matches Haaknat's DNA) and H_in_car (the event that Haaknat got into the car). Note that this graph looks similar to the argument in Fig. 3.4 (p. 40), only with the arrows turned around. This does not mean, however, that the arrows denote causal relations as in Fig. 3.24 (p. 71); the arrows in the graph denote the conditional dependencies (and independencies) of the variables.[17] For each node in the graph the conditional probabilities of the node given its parents can be defined. For example, the conditional probabilities $P(w_1 \mid \text{H\_car})$, $P(\neg w_1 \mid \text{H\_car})$, $P(w_1 \mid \neg(\text{H\_car}))$ and

[16]Pioneering work on Bayesian Networks was done by Pearl. See, for example (Pearl, 1988b, 2000).

[17]Causal Bayesian Networks, in which the arrows denote causal relations, are possible with some additional semantics.

$P(\neg w_1 \mid \neg(H\_car))$ are associated with the node $w_1$. These probabilities constitute the quantitative part of the network and they can be used by an algorithm, which exploits certain independencies that are inherent to Bayesian Networks, to compute the probability of each node in the network.

As was discussed in the previous section, there is a "number problem" when reasoning with probabilities; one way of using conditional probabilities in a meaningful way without committing to specific probabilities is to perform *sensitivity analysis*. With a sensitivity analysis it can be determined what the effects of a change in input are on the output of a probabilistic model. Probabilities still have to be guessed, but the aim of sensitivity analysis is not to calculate the (correct) conditional probability that some hypothesis is the case. Rather, with sensitivity analysis we want to analyse what the effect on the (probability of the) hypothesis is when we change other probabilities that might influence the hypothesis. In other words, we want to know how sensitive the network is to changes in probabilities. The simplest way of performing such an analysis is to systematically change one probability while keeping all others fixed, so that the effect of that value on the rest of the probabilities can be determined. In this way, it can be determined what the probative force of a certain piece of evidence on a hypothesis is.

As an example, consider the small network above. One crucial probability here is $P(w_1 \mid \neg(H\_car))$, that is, how likely is it that the witness saw someone who merely looked like Haaknat? This is a hard probability to judge, as it involves quantifying how many other persons that look like Haaknat can be expected to be in the vicinity of the supermarket at any time. If the witness is assumed to be truthful (i.e. $P(w_1{*} \mid w_1)$ and $P(\neg w_1^* \mid \neg w_1)$ are high), this judgement has a high impact on the probability of the conclusion whether it was Haaknat that got into the car. However, if we now assume that the influence of $w_1^*$ on $w_1$ is slight (i.e. the witness cannot or will not properly tell the difference between Haaknat and someone who looks like Haaknat), then the height of the probability $P(w_1 \mid \neg(H\_car))$ has little influence on the conclusion H_in_car. As Kadane and Schum would put it,[18] if $w_1$ could not tell the difference between Haaknat and someone else, it makes no difference how unlikely it was that someone who only looked like Haaknat got into the car.

The above example of a sensitivity analysis does not require computations in the Bayesian Network: using common sense, we can see that changing the conditional probabilities concerning the link between $w_1^*$ and $w_1$ have an effect on H_in_car. However, in a complex case there will be a large number of propositions that influence each other and more complex sensitivity analyses involve simultaneously varying multiple probabilities. In such cases, a software program that implements a Bayesian Network[19] is an invaluable tool as it keeps track of

[18]Kadane and Schum use a similar example in which a witness testified he saw someone who looked like Sacco.

[19]Examples of such software are Hugin (http://www.hugin.com; accessed on 29th April 2009) and the freely available Genie (http://genie.sis.pitt.edu/; accessed on 29th April 2009).

the probabilistic dependencies between all the propositions and it can automatically perform complex sensitivity analyses through a variety of algorithms. Using these programs, the influence of multiple pieces of evidence on a hypothesis can be established and the possible effects of wrongly-guessed probabilities on the output of a system can be assessed. Kadane and Schum have modelled the combination of the four Wigmore charts concerning the firearm evidence (about which there was disagreement between Kaiser, Kadane and Schum) as a Bayesian Network (Kadane and Schum, 1996, p. 217) and shown that slight changes in some probabilities result in a significant change in the likelihood ratio of some ultimate probandum whilst larger changes in other probabilities do not have a significant effect on the ultimate probandum.

The qualitative part of a Bayesian Network (i.e. the dependency graph) can be used to visualize evidence and inferences from evidence in a way similar to Wigmore graphs and the approach pursued in this book (particularly Chapter 3). Kadane and Schum's work in this respect was already mentioned; other examples are (Levitt and Laskey, 2000; Thagard, 2004). Hepler et al. (2007) recognize this use of Bayesian Networks and propose a way of modelling complex cases using what they call *object-oriented Bayesian Networks*. These networks facilitate the construction of large networks because they are constructed not as a whole, but rather as a collection of smaller networks (or modules) which are placed in a higher level network once approved.[20] As an example, consider the top-level network in Fig. 7.7. The proposition Haaknat guilty at the top of the network is influenced by two nodes. Each of these represents a separate, more detailed network. For example, the node Haaknat near supermarket can be expanded to the network in Fig. 7.8. If Haaknat got into the car he was near the supermarket at the time of the robbery and there may also be other possible witnesses who saw Haaknat before or just after the robbery. Hepler and colleagues discuss more such examples and model part of the Sacco and Vanzetti case as an object-oriented network. They also discuss generic modules which can be re-used, such as modules for the credibility of a witness. Each link between an event and a testimony contains a note signifying a credibility module, viz. Fig. 7.9.

This credibility module is a generic module of the form given in Fig. 7.10. Such generic modules can be used in the construction of complex Bayesian Networks. In the above example, $w_1$'s credibility can be replaced with the credibility module instantiated for $w_1$. Thus Hepler and colleagues' Object-Oriented Bayesian Networks model reasoning and visualization of hierarchical networks which can be refined when desired.

[20]This is similar to the idea of *scenario fragments* as proposed by (Keppens and Schafer, 2006).

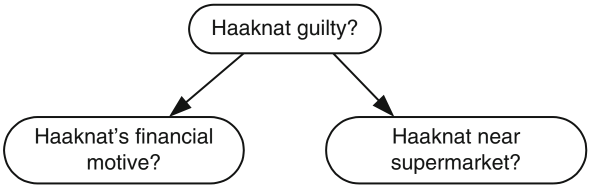



**Fig. 7.7** Top-level Bayesian network for the Haaknat case

**Fig. 7.8** Expansion of the Haaknat near supermarket module from Fig. 7.7

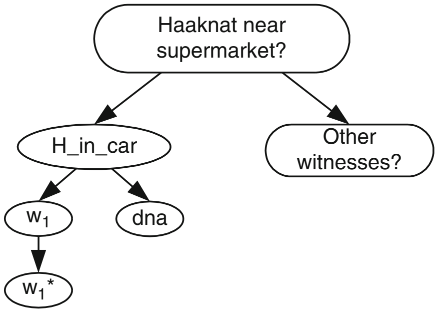
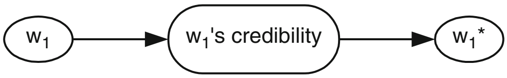



**Fig. 7.9** The credibility attribute in a Bayesian network



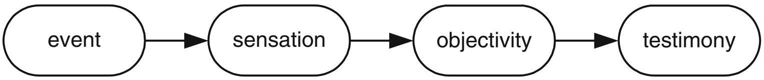




**Fig. 7.10** A generic credibility module

### *7.5.3 Evaluation of Bayesian Approaches*

When reasoning with criminal evidence in the process of proof, we cannot be absolutely certain that a certain hypothesis is true. Investigators have to pursue the hypothesis or hypotheses which they think is *most likely* and a jury in a common law system can convict when it is proven *beyond a reasonable doubt* that the suspect committed the crime. In this way, investigators and decision makers have to make a decision about the probability of hypotheses and the probability that a certain hypothesis is true can increase given evidential data for the hypothesis. Bayesian probability theory is a sound mathematical theory that allows for such reasoning under uncertainty. Bayesian Networks and the software that implements them can be a useful tool in reasoning with evidence. The evidence and the hypotheses and the reasoning in a case can be visualized. Furthermore, Bayesian Networks allow for a relatively easy way of performing useful analyses of the evidence such as sensitivity analysis.

However, the Bayesian approach suffers from a significant drawback, namely the “number problem”: how do we determine the probabilities? In Section 3.1.2 on attacking arguments it was argued that assigning a specific probability to some (conditional) generalization is hard and that an approach which uses (defeasible) preferences is perhaps more natural. Even if techniques are used so that the prior probabilities do not have to be estimated, assessing the likelihood ratio of a probandum given a complex set of evidence is wholly subjective and the harder the assessment, the more people will disagree, as Kadane, Schum and Kaiser have shown.

The “number problem” is obviated if we only use assumed probabilities for sensitivity analyses. With such analyses, the probative force of evidence and the interactions between the various relations that connect the hypotheses and the evidence can be assessed. However, if automated software for Bayesian Networks is used to perform sensitivity analysis, the Bayesian approach suffers from the same “black box” problem as Thagard’s connectionist model. Because of the complexity of the algorithm and Bayesian probability theory in general the results of a multivalued sensitivity analysis can be hard to interpret. This complexity of the Bayesian method makes it less natural for people with no formal training and thus less useful in a sense-making context. Hepler and colleagues argue that the qualitative part of Bayesian networks can be useful for sense-making because it allows for the visualization of complex inference networks. However, care has to be taken with the interpretation of these networks. They do not represent evidential relations (like a Wigmore graph), causal relations (like logical approaches to model-based reasoning) or coherence relations (like in Thagard’s and Josephson’s approaches). Rather, the relations in a Bayesian Network denote dependency and independency relations. Such dependency relations are not as naturally interpretable as the evidential, causal or explain relations used in the present approach. Furthermore, given a Bayesian Network there are implicit independency assumptions which are not apparent from the structure of the network.

To summarize, Bayesian (Network) approaches provide a rationally and computationally sound theory for reasoning with evidence. However, the theory fares less well on naturalness and is thus less suitable for sense-making. A problem of a purely Bayesian theory is that, as in the other formal approaches discussed above, it is impossible to reason *about* the network. However, Nielsen and Parsons (2006) have proposed a novel approach to combining Bayesian Networks through a deliberative dialogue game between autonomous agents. Hepler, Dawid and Leucari present interesting ways of reasoning with graphs and their hierarchical modules bear similarities with aspects of the hybrid theory. For example, generic modules can be likened to story schemes or generalizations. Replacing a single node with a more complex network is similar to evidential and causal refining (see Figs. 3.13 and 3.18). While these ideas do not make the Bayesian approach more natural, they can be used in any visually oriented sense-making tool.

# Chapter 8
# Conclusions

The main aim of this book has been to develop a natural and rationally well-founded theory of reasoning with the evidence and facts in the context of criminal cases. Three important questions in this regard are: What are the roles of arguments and stories in such a theory, what are the necessary elements of such a theory and how can a standard of rationality be set or at least guidelines for rationality be provided? Answers to these research questions have been given in this book. A hybrid theory of stories and arguments that sets a standard of procedural rationality has been developed and presented in a visual, semi-formal style. The hybrid theory has also been formalized; the combination of arguments and stories is modelled as a formal theory, which has specifically been developed with the conceptual, cognitive and computational aims of sense-making in mind. Furthermore, a formal dialogue game for inquiry, which aims to concretise the standard of procedural rationality, has been presented. Finally, a case study has been performed which shows that the hybrid theory is flexible enough to deal with real and large criminal cases. Below the book will be summarized and the results will be discussed.

## 8.1 Summary

Chapter 1 introduces the problem of rational commonsense reasoning with evidence in a criminal legal context. The rational "pursuit of the truth", which is advocated by Wigmore among others, is a process involving masses of evidence and complex reasoning. It is therefore easy to make mistakes common to all types of complex reasoning such as only searching evidence for one hypothesis whilst disregarding alternatives. A theory of proof can aid investigators and decision makers in collecting, organizing and assessing the evidence and the corresponding hypotheses and can also provide a solid basis for a sense-making system such as AVERS. Furthermore, it can encourage reaoners to adhere to a general standard of rational reasoning. The theory should be natural, so that users with little or no formal training can use it, rationally well-founded, so that it is in agreement with the prevailing theories of rational reasoning with evidence, and also should be formally specific to facilitate the design of sense-making software.

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92,
DOI 10.1007/978-94-007-0140-3_8, 

Chapter 2 provides necessary background by giving a general introduction to reasoning with criminal evidence. It introduces the *process of proof* as a general, abstract account of constructing, testing and justifying complex hypotheses. This process of proof lies at the basis of police investigation and judicial decision making in inquisitorial as well as adversarial contexts. The reasoning in the process of proof is a form of inference to the best explanation, where possible hypotheses are first abductively inferred and then subjected to a series of tests which are aimed at finding the best hypothesis. In this process of inferring the best explanation, basic causal knowledge and more complex knowledge about motives and actions and crimes in general is used to explain and predict events. The reasoning the process of proof is *defeasible*. For example, a witness statement that is at first believed may later turn out to be false and a hypothesis that initially looks plausible may be trumped by a hypothesis that better explains the evidence.

There are two important types of knowledge in the process of proof, namely *evidence* and *general commonsense knowledge*. Evidential data, evidence in its most basic form, starts the process of proof by providing initial clues on the basis of which hypothesis can be constructed, steers the testing phase of the process of proof by allowing contradicted hypotheses to be discarded and provides the truths on the basis of which a decision for a particular hypothesis can be justified. General world knowledge, in the form of generalizations and story schemes (i.e. general, abstract scenarios) can be used as the basis for initial hypotheses and to support parts of hypotheses for which there is no direct evidence. In any complex criminal case there are parts of a hypothesis for which no direct supporting evidence can be found and hence some parts of the hypotheses need to be based on general world knowledge, which is part of our (shared) stock of knowledge. However, care should be taken not to accept hypotheses which are only sparsely based on evidence or hypotheses which are based on false commonsense knowledge about the world.

Chapter 3 discusses two main trends in theories for reasoning with criminal evidence: argument-based reasoning and story-based reasoning. Arguments are constructed by performing consecutive reasoning steps, starting with a piece of evidence and reasoning towards some conclusion. Each of these reasoning steps has an underlying *evidential generalization* of the form "$e$ is evidence for $p$" that justifies the step from premises to conclusion and reasoning with arguments can therefore be characterized as *evidential reasoning*. Reasoning with arguments is *dialectical*, in that not only arguments for a conclusion but also arguments against the conclusion and against other arguments are considered. Finally, argumentative reasoning has been called *atomistic* because the various elements of a case (i.e. hypotheses, evidential data) are considered separately and the case is not considered "as a whole". The argument-based approach provides a rationally sound way of analysing and assessing reasoning with evidence, which builds on a significant academic tradition of research on informal and formal argumentation. The argument-based approach allows for a thorough analysis of the individual pieces of evidential data in a case: generalizations used in the reasoning can be supported and attacked in multiple ways and the atomistic nature of arguments means that they are well-suited for exposing sources of doubt in a chain of reasoning. However, this atomistic nature makes the argument-based approach unsuitable for giving an easy overview of the various hypotheses about what happened in the case. Furthermore, not all aspects of causal reasoning such as the prediction of unexpected effects can be found in the purely evidential argument-based approach. Finally, while the idea of evidential arguments and generalizations is based on intuitive and natural concepts, the true usefulness of constructing arguments in organizing evidence has yet to be shown through empirical research.

The story-based approach involves constructing hypothetical stories based on the evidence about what (might have) happened in a case. In this approach one performs *causal reasoning* as the relations between the various events in a story can be expressed as *causal generalizations* of the form "*c* is a cause for *e*". This approach also has a dialectical component in that alternative stories about what happened before, during and after the crime are compared according to their internal coherence and the amount of evidential data they cover. The story-based approach has also been called *holistic* (as opposed to atomistic), because the various elements in the case (i.e. hypotheses, evidential data) are considered as a whole and the individual elements receive less attention.

Furthermore, stories can be considered at various levels of abstraction. The story-based approach has been empirically investigated and is appreciated for its natural account of evidence and scenario's in a criminal case. Stories allow for an easy overview of the case facts, and holistic structures such as story schemes allow for quick construction of hypothetical crime scenarios. The causal nature of stories can be used to further develop a story by explaining and predicting events. However, in the story-based approach the evidential data does not have as clear a place as in the argument-based approach, and therefore its credibility and relevance for the various elements of the hypotheses in a case cannot be checked as easily as with the argumentative approach. Furthermore, a purely story-based approach does not allow for reasoning about a story, that is, it does not allow for reasoning about its plausibility or the extent to which it conforms to the evidence.

From the findings concerning the two separate approaches to reasoning with evidence it is concluded that the best way of reasoning with criminal evidence and crime scenarios is to combine stories and arguments into a *hybrid theory*. Both stories and arguments are necessary for a natural and rationally well-founded theory of reasoning with criminal evidence; stories and arguments are in a sense communicating vessels because in some cases, stories are a natural way of reasoning and in other cases arguments can be used more easily. The hybrid theory has been informally presented in Chapter 4 and formally in Chapter 5. The theory has the full expressiveness of the separate approaches and at the same time adds features that can only exist in a combined theory. The basic idea of the hybrid theory is that hypothetical crime stories can be used to causally explain the explananda (facts to be explained) and that these stories can be supported and contradicted by evidential arguments based on evidence or general world knowledge. In this way, in the hybrid theory it is not only possible to reason *with* stories but also *about* stories: generalizations expressing causal links in a story can be questioned and story schemes that lend plausibility to a story can also be the subject of an argument-based discussion;

thus a consensus about the shared world knowledge that underlies a story can be reached.

Arguments also allow for more precise definitions of the various criteria that can be used to compare explanatory stories. The *evidential support* and *evidential contradiction* of a story is the extent to which a story is supported or contradicted by arguments based on evidence, respectively. Events in a story about which there is no evidence, *evidential gaps*, weaken the story but also point to possible further avenues of investigation. *Plausibility* concerns the plausibility of elements in a story which are not supported by evidence: arguments based on general world knowledge can be used to discuss the (implicit) causal generalizations in the story. Note that this plausibility can be established independently from the evidence; after all, a highly implausible story may be true (i.e. supported by evidence) whilst a highly plausible story may be false (i.e. unsupported or contradicted by evidence). The *completeness* of a story concerns the extent to which it conforms to a plausible story scheme, that is, the extent to which it conforms to our expectations of such a particular situation. Finally, a story should be *consistent* in that it does not contain internal contradictions

Definitions of the various criteria for the quality of stories have been made explicit in Chapter 5. They allow for a detailed and concrete comparison of complex hypotheses (i.e. stories and their corresponding evidence) in the process of proof. The criteria can be phrased as *critical questions* which can be asked for a particular hypothetical story during the process of proof (e.g. "How much and which of the available evidence supports the current hypothesis?"). Given the current procedural conception of rationality, asking these questions in a properly defined dialogue about proof ensures that the process of proof is a rational process in which we can expect that pitfalls such as confirmation bias or unclear justifications of hypotheses are avoided.

A protocol for a proper dialogue about proof is given in Chapter 5. First, the formal hybrid theory is presented as a combination of formal defeasible argumentation and abductive model-based reasoning. After this, a formal dialogue game is defined which models an inquiry dialogue in the context of criminal legal evidence. The dialogue game takes an adversarial approach to inquiry and is constrained by a protocol that ensures relevance. The adversarial approach is modelled by assuming several players who each want to propose, support and defend their own explanatory story whilst simultaneously attacking and criticizing the other players' explanations. In this way, the dialogue game is essentially a model of a critical discussion between investigators or judges. The protocol ensures that all moves are relevant to the current discussion. For example, arguments may only be moved if they improve a player's own explanation or if they worsen an opponent's explanation and only stories that are somehow connected to the explananda may be told.

Chapter 6 presents an analysis of an actual case in the hybrid theory. This procedural and dialectical analysis shows how various alternative stories about what happened in a complex case can be constructed, criticized and compared using evidential arguments. It discusses how the various criteria that determine the quality of a story can be used to guide such an analysis and to uncover sources of doubt in the stories. Furthermore, it contains a detailed discussion on argumentative reasoning about the statements of a single witness (Section 6.4.3).

## 8.2 Results

In this section the results of the current research are discussed. Section 8.2.1 discusses the results important for reasoning with criminal evidence in general and Section 8.2.2 discusses the specific results of the formal model.

### *8.2.1 Reasoning with Criminal Evidence*

The current hybrid theory's contribution lies in particular in its full integration of arguments and stories into one theory for reasoning with criminal evidence. Whilst Modified Wigmorean Analysis and the Anchored Narratives Theory both allow the use of stories as well as arguments (or at least some sort of argumentative inference), neither of these theories in my view fully integrates the two. The hybrid theory bridges the gap between the Modified Wigmorean Analysis and the Anchored Narratives Theory. Proponents of the first approach have argued that reasoning with (argument) charts is the best rational way of reasoning with evidence. They argue that stories are mainly useful for organizing and presenting the evidence and that they play a largely psychological (as opposed to a critically rational) role in evidence evaluation. In this book I have shown that, when combined with arguments and open to criticism, causal stories in inference to the best explanation are not only a natural but also a rationally way of reasoning. Proponents of the more story-centred approaches have argued that the only natural way in which people reason with evidence is through stories and that all reasoning in a case takes the form of a story. However, often reasoning from evidential data to some conclusion takes the form of a syllogistic argument, and the hybrid theory shows how stories can be grounded in evidence through such arguments.

The discussion of the separate argument-based and story-based approaches has also shown that stories and arguments are both are necessary for a natural and rationally well-founded theory of reasoning with criminal evidence. By retaining the advantages and flexibility of the separate approaches whilst at the same time solving some of their problems by integrating the two approaches, the hybrid theory acknowledges this interaction between evidence, arguments and stories. Holistic stories provide an overview of the various possible scenarios in a case and can be used to make sense of a complex mass of evidence. They can be used in a relatively simple way to construct complex new hypotheses in the discovery phase of the process of proof. Furthermore, when combined with arguments, stories are useful when justifying one's decision in a complex case because they help make sense of the evidence and the events that can be inferred from the evidence. Thus, a decision can be more easily checked by third parties. Arguments provide a natural connection between the evidential data and the facts of the case. Notions such as critical questions and attackers can be used to reason about the relevance and inferential force of an argument and the focused and atomistic way of reasoning with arguments allows for a detailed analysis of the individual pieces of evidence, the hypothetical stories and the commonsense world knowledge that is used in the reasoning. The direct interaction between arguments and stories, that is, the ways in which arguments can directly support and contradict stories and the ways in which the criteria for judging a story's quality are dependent on evidential arguments, are unique to the hybrid theory and allow for a natural and rationally well-founded theory of inference to the best explanation.

In a sense, stories and arguments in the hybrid theory act as communicating vessels. At some point in a case, only one individual state or event will be at issue and in such cases it is most natural to reason with evidential arguments. For example, we might want to know if a suspect was at a particular location at a particular time and consider all the evidence for and against this fact. However, at other points in an investigation, we might use a story to fill an evidential gap and see whether assuming the gap-filling event still allows for a plausible story. For example, if no evidence for the suspects location at a particular important time is available we can assume that the suspect actually was at the location and, given other evidence about the suspect's whereabouts at other points in time, see if the causal and temporal structure of the story is still plausible.

In addition to the improvements gained by combining stories and arguments, the current work has also improved on the existing separate argument-based and story-based accounts of separately reasoning with stories or arguments. The current work[1] improves on the argument-based work in evidence theory, most notably Modified Wigmorean Analysis' chart method, on several points. The detailed formal analysis of reasoning with different types of evidence has yielded standard argumentation schemes or prima facie reasons for particular types of evidence, which can easily be used in the dialectical process argumentation. This formal analysis has also provided a systematic account of (conditional and non-conditional) generalizations and various ways to attack these generalizations. In addition to aiding in the detailed and systematic analysis of reasoning with evidence, the formalization of arguments also provides a solid logical underpinning for the argument-based component of the hybrid theory. For example, the difference between reasoning about the validity of generalizations and exceptions to generalizations has been logically modelled and the argumentation-based semantics that governs the status of arguments provides a solid mathematical basis to the comparison of arguments.

Reasoning with stories has also been elaborated upon in this book. Clarity has been provided regarding the rational (as opposed to the psychological) role of stories in reasoning with evidence, namely that they serve as the explanations for the explananda in the process of inference to the best explanation. Furthermore, the

[1] Which is based on earlier research by Bex et al. (2003).

ways in which a story's quality can and should be determined and tested have also been expanded, as the criteria that play a role in determining the quality of a story are also extensively discussed and precisely defined. Particularly the definitions for the coherence of stories are an improvement over existing work in legal psychology: Bennett and Feldman do not further define their notion of "ambiguous connections" and Pennington and Hastie do not define the notion of plausibility and only partly define completeness. Both these notions (i.e. plausibility and completeness) have been defined precisely in the current work, and the distinction between the plausibility of a story (i.e. the plausibility of the generalizations expressing the causal relations) and the plausibility of the evidential generalizations that warrant the argumentative inferences from evidence to an event has proven to be useful. It allows for a discussion of the plausibility of a story independently of the evidential data and refines the notion of anchoring as proposed by Crombag, Van Koppen and Wagenaar. The discussion of story schemes further explores Pennington and Hastie's and Bennett and Feldman's basic ideas. With regards to a story's completeness, the new distinction between completing a story scheme (every element of the scheme has a corresponding event in the story) and fitting a story scheme (every event in the story has a corresponding event in the scheme) ensures not only that a story has all its parts but also that there are no loose ends. Additionally, further uses of story schemes (e.g. as templates in abductive reasoning) have also been briefly mentioned and are possible in the current hybrid theory and dialogue game. In this way, influential work in AI and cognitive psychology (e.g. Rumelhart, 1975, Schank and Abelson, 1977) has been given a clear place in the hybrid theory of reasoning with stories and criminal evidence.

The procedural view on rationality underlying the present approach, in which the dialectical processes of argument and counterargument and explanation and alternative explanation important, is explicit in the underlying theories of dialectical argumentation and inference to the best explanation. The various ways of attacking arguments and generalizations and the critical questions associated with argumentation schemes (and the associated undercutters for prima facie reasons) serve as guidelines for the dialectical process with arguments. Similarly, the ways in which stories can be attacked with arguments and the critical questions for stories provide guidelines for the dialectical process of inference to the best explanation. Negative answers to critical questions can be seen as pitfalls which need to be avoided when reasoning with arguments and stories in a criminal case. The formal protocol supplies further rules for a rational discussion of the evidence and the facts. Given the current procedural conception of rationality, asking critical questions in a properly defined dialogue about proof ensures that the process of proof is a rational process in which we can expect that pitfalls such as confirmation bias or unclear justifications of hypotheses are avoided.

The findings discussed in this section are largely supported by the Anjum case study. In this case study, the different stories about who killed Leo represent natural explanations for his death. These stories are grounded in the evidence and can be rationally criticized using arguments based on evidence and general knowledge. The constant interaction between causal and evidential reasoning is clearly shown as both evidential and causal information play important roles in the case. The procedural character of the analysis allows for an incremental way of analysing a case, which benefits the comprehensibility and readability of the analysis. This readability is further improved by the alternative stories that effectively summarize the various views in the case. These stories can be criticized using the formal criteria (e.g. evidential support, plausibility, completeness) and the evidential gaps are marked so the danger of a good story pushing out a bad story is avoided.

### *8.2.2 Formal Theories of Defeasible Reasoning*

In addition to providing answers for evidence theorists and legal psychologists, this book also set out to advance the multidisciplinary field of Artificial Intelligence and Law and in particular the side of AI and Law that concerns defeasible reasoning in the context of the law.

The theory as developed in this book has been expressly developed to underpin a sense-making tool. Hence, it is one of the few logical theories which has conceptual, cognitive as well as computational aims. Furthermore, while other sense-making tools (e.g. Verheij's (1999) ArguMed) also have a logical underpinning, their logic was not developed especially with sense making for the process of proof in mind. In other cases (e.g. Araucaria, AIF), first a more informal sense-making tool was developed and only then has the connection between the tool and more formal logics been explored. In the current project, the logic was developed concurrently with the prototype system AVERS. The logical model of causal stories combined with evidential arguments is directly implemented in the system; the results of the tests performed with the system and the informal contact with various teams of the Dutch police force and Dutch police academy have influenced the theory and strengthened the claim that the hybrid theory is close to how actual reasoning with evidence in an investigation context is performed. Furthermore, the case study in Chapter 6 supports these findings and is at the same time one of the few thorough and large case studies of both a defeasible logic for argumentation and causal model-based techniques.

The design of the hybrid theory, which as mentioned above has conceptual, cognitive and computational aims, provides new insights which are interesting for formal defeasible reasoning in general and logical inference to the best explanation in particular. In real domains such as criminal investigation, the logical or mathematical modelling of inference to the best explanation is a hard enterprise. One reason for this is that a logical or mathematical theory which is used to model the explanations in the domain is itself complex. Hence, a model of the domain in a logical theory is too complex to be constructed or understood by (logical or mathematical) laymen. For example, a proper Bayesian Network cannot be constructed or fully appreciated without in depth knowledge of the probabilities and dependencies expressed in the mathematical model underlying the network. In such a case, the user of a model or system based on a logical model does not know if his or her beliefs about the case are correctly expressed.

The hybrid theory is based on a logical model that models explanations as stories, causal networks in which events are connected with simple, case-specific generalizations. The theory is based on research on how we understand the world around us and how the human memory is organized; stories and story schemes are natural tools we use in our everyday reasoning about a complex world. Therefore, no formal mathematical training is needed and models of relatively complex cases such as the Anjum murder case can be constructed and understood by laymen. A formal model of argumentation, based on the intuitive concepts of argument and counterargument, is used to reason about the formal causal model (i.e. a hypothetical story). This combination of stories and critical argumentation makes the hybrid theory ideal for sense-making: the knowledge represented in the system (i.e. the stock of knowledge and the hypothetical stories) can be constructed incrementally and is defeasible in that knowledge which is accepted at one point may be rejected at a later time when new evidence that contradicts this knowledge becomes available. Thus the users of the system can try to reach a cognitive consensus about the model of a case that is compatible with the evidence in a natural and rational way.

A further problem of using a complex logical theory as the basis for a (sense-making or automated reasoning) system that performs or aids in inference to the best explanation is that any output generated by the system (e.g. which explanation is the best) is hard to interpret or hard to verify because the process through which the system arrived at this output is not transparent. The current criteria for the quality of a story, however, are based on features of stories that have been proposed and tested in theories of story understanding, such as causal connectivity and adherence to some general story scheme. The output given by the intuitive argument-based semantics in combination with the ordering on explanations is easy to understand and it can be easily checked how the various components of a particular case interact and what can be done to change, for example, the ordering on explanations. The adversarial dialogue game provides a procedural standard of rationality by always allowing a player to improve his own explanation and by ensuring that more than one explanation is put forward.

The formal hybrid theory is the first to allow for the use of formal argumentation to construct and talk about some (causal) model. In this way, it improves on traditional model-based modelling techniques by allowing for an explicit discussion of the plausibility of the causal model. Furthermore, whilst hierarchies of causal explanations have been proposed in other work, the use of such hierarchies (i.e. stories that correspond to some abstract story schemes) in determining the best explanation is new. The dialogue game is one of the first formal dialogue games that model an inquiry dialogue. Furthermore, its precise form is to my knowledge new: Nielsen and Parsons' dialogue game, which allows one to build a Bayesian network, is to my knowledge one of the few other games that allow the reasoner to build a non-argumentative model.

The insights provided by formalizing the hybrid theory and the dialogue game for inquiry are not just useful in a logical theory for crime investigation, as any complex domain involving either evidence or some sort of model-based reasoning can benefit from these insights. The formal criteria can also be applied to other models that use (causal) models for commonsense reasoning and the possibility of constructing and discussing a model in a formal dialogue game using arguments ensures that the model meets a certain standard of rationality.

## 8.3 Suggested Topics for Further Research

In this book, I have aimed at proposing a complete theory of rationally reasoning with criminal evidence. However, there are multiple avenues of research by which the informal as well the formal hybrid theory might be improved.

As was argued at the beginning of this book, the process of proof is not reserved for the context of criminal investigation or criminal trial. This does not mean, however, that there are no specific features of the process of proof which are specific to the criminal legal context. While in Section 2.2.1, the relation of the general process of proof to the process of investigation and trial was very briefly discussed, more research should be done to show what the implications of the current theory for the criminal legal domain are. In the current research I have tried to (implicitly) retain the connection to the legal and investigative practice but more can be learnt from jurisprudence and the writings on evidence law and intelligence analysis, which represent practical knowledge gained by people who actively work with the process of proof every day.

The (empirical) research on sense-making by Susan van den Braak and colleagues (2007, 2008, 2010) can play an important role in the connection between the current research and investigative practice. This Utrecht-based part of the current project implements a version of the formal-logical theory in AVERS[2] and tests it with police intelligence analysts and students of the Dutch police academy. This can provide interesting results as to the naturalness of the hybrid theory and can point to improvements of the hybrid theory. Further developing and testing AVERS – for example, by implementing the dialogue game in some way – is needed and thus provides for interesting future research.

The formal theory presented in Chapter 5 presents a solid basis for the conceptual (informal) hybrid theory of arguments and stories. However, by itself the theory is not complete or without weaknesses. The theory can be enriched in various ways. For example, time and temporal reasoning can be added and the various ways of predictive reasoning can be further expanded in the dialogue game. A relevant issue is the discussion (and proof) of formal properties of the theory, which could answer some important questions concerning the rationality and correctness of the hybrid

[2]The last version of AVERS works with a slightly adapted version of the "static" hybrid theory. The dialogue game has not been explicitly implemented in the system.

theory.[3] A further interesting question regarding the formal dialogue game is to see how changes in player strategies, winning conditions and turntaking influence the outcome of the dialogue. In this way, heuristics in the form of dialogue games can be developed for various purposes, such as guiding police investigation or helping the prosecutor or the defence counsel prepare a case.

[3]Examples of such questions are "can we be sure that the dialogue game always leads to a correct formal theory *AET*?" and "does a story that completes a scheme always have all of its parts?".

# References

Allen, J.F. and Ferguson, G. (1994) Actions and events in interval temporal logic. *Journal of Logic and Computation* 4:5, 531–579.

Amgoud, L. and Cayrol, C. (2002) A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence* 34:1, 97–215.

Anderson, T.J. (1999) On generalizations I: A preliminary exploration. *South Texas Law Review* 40, p. 455.

Anderson, T.J., Schum, D.A. and Twining, W.L. (2005) *Analysis of Evidence*, 2nd edition, Cambridge University Press, Cambridge.

Aristotle. (2005) Poetics. In Butcher, S.H. (eds.), *Wordplay Archives*, http://wordplayer.com/archives/poetics.txt

Ashley, K.D. (1991) *Modeling Legal Arguments: Reasoning with Cases and Hypotheticals*, MIT Press, Cambridge (Massachusetts).

Atkinson, K., Bench-Capon, T. and McBurney, P. (2006) Computational representation of practical argument. *Synthese* 152:2, 157–206.

Bal, M. (1985) *Narratology: Introduction to the Theory of Narrative*, University of Toronto Press, Toronto.

Bayes, R.T. (1763) An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society London* 53, 370–418.

Bench-Capon, T.J.M. (2003) Persuasion in practical argument using value – based argumentation frameworks. *Journal of Logic and Computation* 13:3, 429–448.

Bennett, W.L. and Feldman, M.S. (1981) *Reconstructing Reality in the Courtroom: Justice and Judgment in American Culture*, Methuen – Tavistock, London.

Bex, F.J., van Koppen, P.J., Prakken, H. and Verheij, B. (2010) Evidence for a good story – a hybrid theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law* 18, 2.

Bex, F.J. (2009) Analysing stories using schemes. In Kaptein, H., Prakken, H. and Verheij B. (eds.), *Legal Evidence and Proof: Statistics, Stories, Logic*, 93–116, Ashgate Publishing, Aldershot.

Bex, F.J., Bench-Capon, T. and Atkinson, K. (2009) Did he jump or was he pushed? abductive practical reasoning. *Artificial Intelligence and Law* 17:2, 79–99.

Bex, F.J., Braak, S.W. van den, Oostendorp, H. van, Prakken, H., Verheij, B. and Vreeswijk, G. (2007b) Sense – making software for crime investigation: how to combine stories and arguments? *Law, Probability and Risk* 6, 145–168.

Bex, F.J. and Prakken, H (2004) Reinterpreting arguments in dialogue: an application to evidential reasoning. *JURIX 2004: The 17th Annual Conference*, 119–129, IOS Press, Amsterdam.

Bex, F.J. and Prakken, H. (2008) Investigating stories in a formal dialogue Game. *Computational Models of Argument. Proceedings of COMMA 2008*, 73–84, IOS Press, Amsterdam.

Bex, F.J., Prakken, H., Reed, C. and Walton, D.N. (2003) Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 11, 125–165.

F. Bex, *Arguments, Stories and Criminal Evidence*, Law and Philosophy Library 92, DOI 10.1007/978-94-007-0140-3, © Springer Science+Business Media B.V. 2011

Bex, F.J., Prakken, H. and Verheij, B. (2006) Anchored narratives in reasoning about evidence. *JURIX 2006: The 19th Annual Conference*, 11–20, IOS Press, Amsterdam.

Bex, F.J., Prakken, H. and Verheij, B. (2007a) Formalising argumentative story – based analysis of evidence. *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, 1–10, ACM Press, New York, (New York).

Bex, F.J. and Verheij, B. (2009) Het onderbouwen van een feitelijk oordeel in een strafzaak: methode, casus, aanbevelingen (Grounding a judgement about the facts in a criminal case: method, case, recommendations). In Van Koppen, P.J., Merkelbach, H., Jelicic, M., and de Keijser, J.W. (eds.), *Reizen Met Mijn Rechter: Psychologie Van Het Recht*, Kluwer, Deventer.

Black, J.B. and Wilensky, R. (1979) An evaluation of story grammars. *Cognitive Science* 3, 213–230.

Bondarenko, A., Dung, P.M., Kowalski, R.A. and Toni, F. (1997) An abstract, argumentation – theoretic approach to default reasoning. *Artificial Intelligence* 93:1–2, 63–101.

Braak, S.W. van den (2010) *Sensemaking Software for Crime Analysis*. Doctoral dissertation, Intelligent Systems Group, Utrecht University (SIKS Dissertation Series No. 2010–2012).

Braak, S.W. van den, Oostendorp, H. van, Vreeswijk, G. and Prakken, H. (2008) Representing Narrative and Testimonial Knowledge in Sense – making Software for Crime Analysis. *JURIX 2008: The 21st Annual Conference*, 160–169, IOS Press, Amsterdam.

Braak, S.W. van den, Vreeswijk, G. and Prakken, H. (2007) AVERs: An argument visualization tool for representing stories about evidence. *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, ACM, New York, NY (New York).

Braak, S.W. van den, Oostendorp, H. van, Prakken, H. and Vreeswijk, G.A.W. (2006) A critical review of argument visualization tools: do users become better reasoners. *ECAI – 06: Workshop on Computational Models of Natural Argument*. Trento

Burch, R. (2008) Charles sanders peirce, In Zalta, E.N. (eds.), *Stanford Encyclopedia of Philosophy,* Fall 2008 edition, http://plato.stanford.edu/archives/fall2008/entries/peirce/

Bylander, T., Allemang, D., Tanner, M.C. and Josephson, J.R. (1991) The computational complexity of abduction. *Artificial Intelligence* 49:1–3, 25–60.

Chesñevar, C.I., Maguitman, A.G. and Loui, R.P. (2000) Logical models of argument. *ACM Computing Surveys (CSUR)* 32:4, 337–383.

Cohen, L.J. (1977) *The Probable and The Provable*, Oxford University Press, Oxford.

Console, L. and Dupré, D.T. (1994) Abductive reasoning with abstraction axioms. *Lecture Notes in Computer Science,* Vol. 810, 98–112, Springer, Berlin.

Console, L., Dupré, D.T. and Torasso, P. (1991) On the relationship between abduction and deduction. *Journal of Logic and Computation* 1, 661–690.

Console, L. and Torasso, P. (1991) A spectrum of logical definitions of model – based diagnosis. *Computational Intelligence* 7:3, 133–141.

Crombag, H.F.M. and Israëls, H. (2008) *Moord in Anjum – Te veel niet gestelde vragen*, Boom Juridische Uitgevers, Den Haag.

Crombag, H.F.M., Wagenaar, W.A. and Koppen, P.J. van. (1994) *Dubieuze Zaken: De Psychologie van Strafrechtelijk Bewijs*, 2nd edition, Contact, Amsterdam.

Dawid, P. (2005) *Probability and Proof – Appendix to* “*Analysis of Evidence*”. http://www.cambridge.org/catalogue/catalogue.asp?isbn=052167316X&ss=res Last accessed 30–1–2006.

De Poot, C.J., Bokhorst, R.J., Koppen, P.J. van and Muller, E.R. (2004) *Rechercheportret – Over Dillemma’s in de Opsporing*, Kluwer, Alphen a.d. Rijn.

Dennett, D.C. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*, MIT press, Cambridge (Massachusetts).

Dingley, A.M. (1999) The ballpoint case: a wigmorean analysis. In Nijboer, J.F. and Malsch, M. (eds.), *Complex Cases: Perspectives on the Netherlands Criminal Justice System*, 179–188, Thela Thesis, Amsterdam.

Dung, P.M. (1995) On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n – person games. *Artificial Intelligence* 77:2, 321–357.

Eco, U. and Sebeok, T.A. (1983) *The Sign of Three: Dupin, Holmes, Pierce*, Indiana University Press, Bloomington (Indiana).

Eemeren, F.H. van and Grootendorst, R. (2004) *A Systematic Theory of Argumentation: The Pragma – Dialectical Approach*, Cambridge University Press, Cambridge.

Eemeren, F.H. van, Grootendorst, R., Henkemans, F.S., Blair, J.A., Johnson, R.H., Krabbe, E.C.W., Plantin, C., Walton, D.N., Willard, C.A. and Woods, J. (1996) *Fundamentals of Argumentation Theory*, Lawrence Erlbaum, Hillsdale (New Jersey).

Einhorn, H.J. and Hogarth, R.M. (1986) Judging probable cause. *Psychological Bulletin.* 99:1, 3–19.

Feteris, E.T. (1999) What went wrong in the ballpoint case?. In Malsch, M. and Nijboer, J.F. (eds.), *Complex Cases: Perspectives on the Netherlands Criminal Justice System*, 159–179, Thela Thesis, Amsterdam.

Freeman, J.B. (1991) *Dialectics and the Macrostructure of Arguments: A Theory of Argument Structure*, Foris Publications, Berlin.

Gabbay, D.M. and Woods, J. (2006) *Advice on Abductive Logic*, 14th edition, Oxford University Press, Oxford.

Gabbay, D.M., Hogger, C.J. and Robinson, J.A. (1993) *Handbook of Logic in Artificial Intelligence and Logic Programming*, *Volume 1*, *Logical Foundations*, Oxford University Press, Oxford.

Gordon, T.F., Prakken, H. and Walton, D. (2007) The Carneades model of argument and burden of proof. *Artificial Intelligence* 171:10–15, 875–896.

Hage, J.C. (1996) A theory of legal reasoning and a logic to match. *Artificial Intelligence and Law* 4:3, 199–273.

Hage, J.C. and Verheij, B. (1999) The law as a dynamic interconnected system of states of affair: a legal top ontology. *International Journal of Human – Computer Studies* 51, 1043–1077.

Hanson, N.R. (1962) *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*, The Scientific Book Guild, London.

Hepler, A.B., Dawid, A.P. and Leucari, V. (2007) Object – oriented graphical representations of complex patterns of evidence. *Law, Probability and Risk* 6, 275–293.

Heuer, R.J. (1999) *Psychology of Intelligence Analysis*, Center for the Study of Intelligence, Central Intelligence Agency.

Hielkema, J. (1999) Experts in dutch criminal procedure. In Malsch, M. and Nijboer, J.F. (eds.), *Complex Cases: Perspectives on the Netherlands Criminal Justice System*, 27–46, Thela Thesis, Amsterdam.

Hume, D. (1888) A treatise of human nature. In Selby–Bigge, L.A. (eds.), *Hume's Treatise of Human Nature*, Clarendon Press, Oxford.

Johnson, N.S. and Mandler, J.M. (1980) A tale of two structures: Underlying and surface forms in stories. *Poetics* 9, 51–86.

Josephson, J.R. (2002) On the proof dynamics of inference to the best explanation. In MacCrimmon, M. and Tillers, P. (eds.), *The Dynamics of Judicial Proof – Computation, Logic and Common Sense*, 287–306, Physica Verlag, Berlin.

Josephson, J.R. and Josephson, S.G. (1994) *Abductive Inference: Computiation, Philosophy, Technology*, Cambridge University Press, Cambridge.

Kadane, J.B. and Schum, D.A. (1996) *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*, Wiley, New York (New York).

Kautz, H.A. (1991) A formal theory of plan recognition and its implementation. In Allen, J.F., Kautz, H.A., Pelavin, R.N., and Tenenberg, J.D. (eds.), *Reasoning About Plans*, 69–124, Morgan Kaufmann, San Mateo (California).

Keppens, J. and Schafer, B. (2005) Assumption – based peg unification for crime scenario modelling. *JURIX 2005: The 18th Annual Conference*, 49–58, IOS Press, Amsterdam.

Keppens, J. and Schafer, B. (2006) Knowledge based crime scenario modelling. *Expert Systems with Applications* 30, 203–222.

Kerstholt, J.H. and Eikelboom, A.R. (2007) Effects of prior interpretation on situation assessment in crime analysis. *Journal of Behavioral Decision Making* 20:5, 455–465.

Knigge, G. (2001) *Leerstukken Van Het Strafprocesrecht*, 5th edition, Kluwer, Deventer.

Konolige, K. (1994) Using default and causal reasoning in diagnosis. *Annals of Mathematics and Artificial Intelligence* 11:1, 97–135.

Kowalski, R. and Sergot, M. (1986) A logic – based calculus of events. *New Generation Computing* 4:5, 67–94.

Lempert, R. (1986) The new evidence scholarship: analyzing the process of proof. *Boston University Law Review* 66, 439.

Lent, L. van. (2008) *Externe Openbaarheid in Het Strafproces.*, Dissertation Utrecht University, Utrecht.

Levitt, T.S. and Laskey, K.B. (2000) Computational inference for evidential reasoning in support of judicial proof. *Cardozo Law Review* 22, 1691.

Loftus, E.F. (1996) *Eyewitness Testimony*, 2nd edition, Harvard University Press, Cambridge (Massachusetts).

Loui, R.P. (1987) Defeat among arguments: a system of defeasible inference. *Computational Intelligence* 3:1, 100–106.

Loui, R.P. (1995) Hart's critics on defeasible concepts and ascriptivism. *Proceedings of the 4th International Conference on Artificial Intelligence and Law*, 21–30, ACM Press, New York (New York).

Loui, R.P. (1998) Process and policy: resource – bounded nondemonstrative reasoning. *Computational Intelligence* 14:1, 1–38.

Loui, R.P. and Norman, J. (1995) Rationales and argument moves. *Artificial Intelligence and Law* 3:3, 159–189.

Lucas, P. (1997) Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review* 12, 109–146.

MacCrimmon, M., and Tillers, P. (eds.) (2002) *The Dynamics of Judicial Proof: Computation, Logic, and Common Sense*, Physica – Verlag, Berlin.

Malsch, M. and Freckelton, I. (2009) The evaluation of evidence: differences between legal systems. In Kaptein, H., Prakken, H., and Verheij, B. (eds.), *Legal Evidence and Proof: Statistics, Stories, Logic*, 117–134, Ashgate, Aldershot.

Mandler, J.M. and Johnson, N.S. (1977) Remembrance of things parsed: Story structure and recall. *Cognitive Psychology* 9, 111–151.

Modgil, S. (2007) An abstract theory of argumentation that accommodates defeasible reasoning about preferences. *Lecture Notes in Computer Science* 4724: 648–659, Springer, Berlin.

Nielsen, S.H. and Parsons, S. (2006) An application of formal argumentation: Fusing Bayes nets in MAS. *Computational Models of Argument: Proceedings of COMMA 2006*, IOS Press, Amsterdam.

Nijboer, J.F. (1993) Common law tradition in evidence scholarship observed from a continental perspective. *The American Journal of Comparative Law* 51:2, 299–338.

Nijboer, J.F. (2000) *Strafrechtelijk Bewijsrecht*, 4th edition, Ars Aequi Libri, Nijmegen.

Nijboer, J.F. and Sennef, A. (1999) Justification. In Nijboer, J.F. and Malsch, M. (eds.), *Complex Cases: Perspectives on the Netherlands Criminal Justice System*, 11–26, Thela Thesis, Amsterdam.

Pardo, M.S. and Allen, R.J. (2007) Juridical proof and the best explanation. *Law and Philosophy* 27, 223–268. Springer.

Pearl, J. (1988a) Embracing causality in default reasoning. *Artificial Intelligence* 35, 259–271.

Pearl, J. (1988b) *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo (California).

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge.

Peirce, C.S. (1931) Hartshorne, C., Weiss, P., and Burks, A. (eds.), *Collected Papers of Charles Sanders Peirce*, Harvard University Press, Cambridge (Massachusetts).

Pennington, N. and Hastie, R. (1986) Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology* 51, 242–258.

Pennington, N. and Hastie, R. (1988) Explanation – based decision making: effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14:3, 521–533.

Pennington, N. and Hastie, R. (1992) Explaining the evidence: tests of the story model for juror decision making. *Journal of Personality and Social Psychology* 62:2, 189–206.

Pennington, N. and Hastie, R. (1993a) Reasoning in explanation – based decision making. *Cognition* 49:1–2, 123–163.

Pennington, N. and Hastie, R. (1993b) The story model for juror decision making. In Hastie, R. (eds.), *Inside the Juror: The Psychology of Juror Decision Making*, Cambridge University Press, Cambridge.

Pollock, J.L. (1987) Defeasible reasoning. *Cognitive Science* 11:4, 481–518.

Pollock, J.L. (1995) *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, Cambridge (Massachusetts).

Poole, D. (1988) A logical framework for default reasoning. *Artificial Intelligence* 36:1, 27–47.

Poole, D. (1994) Representing diagnosis knowledge. *Annals of Mathematics and Artificial Intelligence* 11:1, 33–50.

Poole, D., Mackworth, A. and Goebel, R. (1997) *Computational Intelligence: A Logical Approach*, Oxford University Press, Oxford.

Prakken, H. (1997) *Logical Tools for Modelling Legal Argument*, Kluwer Academic Publishers, Dordrecht.

Prakken, H. (2005a) A study of accrual of arguments, with applications to evidential reasoning. *Proceedings of the 10th International Conference on Artificial Intelligence and Law*, 85–94, ACM, New York (New York).

Prakken, H. (2005b) Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation* 15, 1009–1040.

Prakken, H. (2006) Formal systems for persuasion dialogue. *The Knowledge Engineering Review* 21:02, 163–188.

Prakken, H. and Sartor, G. (1996) A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law* 4:3, 331–368.

Prakken, H. and Sartor, G. (1997) Argument – based extended logic programming with defeasible priorities. *Journal of Applied Non – classical Logics* 7, 25–75.

Prakken, H. and Vreeswijk, G. (2002) Logics for defeasible argumentation. In Goebel, R. and Guenthner, F. (eds.), *Handbook of Philosophical Logic*, 219–318, Kluwer Academic Publishers, Dordrecht.

Propp, V. (1968) *The Morphology of the Folktale*, University of Texas Press, Austin (Texas).

Quine, W.V.O. and Ulian, J. (1970) *The Web of Belief*, Random House, New York (New York).

Rahwan, I., Ramchurn, S.D., Jennings, N.R., McBurney, P., Parsons, S. and Sonenberg, L. (2004) Argumentation – based negotiation. *The Knowledge Engineering Review* 18:04, 343–375.

Rahwan, I. and Reed, C. (2009) The argument interchange format. In Rahwan, I. and Simari, G. (eds.), *Argumentation in Artificial Intelligence*, Springer.

Reed, C. and Rowe, G. (2004) Araucaria: software for argument diagramming, analysis and representation. *International Journal of AI Tools* 13, 961–980.

Reed, C., Walton, D. and Macagno, F. (2007) Argument diagramming in logic, law and artificial intelligence. *Knowledge Engineering Review* 22:1, pp 87–109.

Reichenbach, H.F.G. (1978) Reichenbach, M. and Cohen, R.S. (eds.), *Selected Writings: 1909–1953*, Springer, Berlin.

Reiter, R. (1980) A logic for default reasoning. *Artificial Intelligence* 13, 81–132.

Rescher, N. (1977) *Dialectics: A Controversy – Oriented Approach to the Theory of Knowledge*, State University of New York Press, Albany (New York).

Rider, Y. and Thomason, N. (2008) Cognitive and pedagogical benefits of argument mapping: L.A.M.P. guides the way to better thinking. In Okada, A., Buckingham Shum, S. and Sherborne, T. (eds.), *Knowledge Cartography: Software Tools and Mapping Techniques*, Springer, London.

Rumelhart, D.E. (1975) Notes on a schema for stories. In Bobrow, D.G. and Collins, A. (eds.), *Representation and Understanding: Studies in Cognitive Science*, Academic Press, New York (New York).

Sartor, G. (2008) Legality Policies and Theories of Legality: From 'Bananas' to Radbruch's Formula. *EUI Working Papers LAW* 27.

Schank, R.C. (1975) The structure of episodes in memory. In D. Bobrow and A. Collins (eds.), *Representation and Understanding: Studies in Cognitive Science*, 235–259, Academic Press, New York.

Schank, R.C. (1986) *Explanations Patterns: Understanding Mechanically and Creatively*, Lawrence Erlbaum, Hillsdale (New Jersey).

Schank, R.C. and Abelson, R.P. (1977) *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*, Lawrence Erlbaum, Hillsdale (New Jersey).

Schum, D.A. (1994) *The Evidential Foundations of Probabilistic Reasoning*, Northwestern University Press, Evanston (Illinois).

Schum, D.A. (2001) Species of abductive reasoning in fact investigation in law. *Cardozo Law Review* 22, 1645–1681.

Schum, D.A. (2005) *Narratives in Intelligence Analysis: Necessary but Often Dangerous*, Evidence Research, London.

Shanahan, M. (1989) Prediction is deduction but explanation is abduction. *Proceedings of International Joint Conference on Artificial Intelligence* 89, 1055–1060.

Shen, Q., Keppens, J., Aitken, C., Schafer, B. and Lee, M. (2007) A scenario – driven decision support system for serious crime investigation. *Law, Probability and Risk* 5:2, 87–117.

Simon, H.A. (1982) *Models of Bounded Rationality*, MIT Press, Cambridge (Massachusetts).

Simon, D. (2001) A third view of the black box: cognitive coherence in legal decision making. *University of Chicago Law Review* 71, 511–586.

Stein, N.L. and Glenn, C.G. (1979) An analysis of story comprehension in elementary school children. *New directions in discourse processing* 2, 53–120.

Thagard, P. (1988) *Computational Philosophy of Science*, MIT Press, Cambridge (Massachusetts).

Thagard, P. (1989) Explanatory coherence. *Behavioral and Brain Sciences* 12:3, 435–502.

Thagard, P. (2004) Causal inference in legal decision making: explanatory coherence vs. bayesian networks. *Applied Artificial Intelligence* 18:3, 231–249.

Thagard, P. (2005) Testimony, credibility, and explanatory coherence. *Erkenntnis* 63:3, 295–316.

Thagard, P. and Shelley, C.P. (1997) Abductive reasoning: logic, visual thinking, and coherence. In Dalla Chiara, M.L. (eds.), *Logic and Scientific methods*, Kluwer Academic Publishers, Dordrecht.

Thorndyke, P.W. (1977) Cognitive structures in comprehension and memory of narrative discours. *Cognitive Psychology* 9, 77–110.

Tillers, P. (2005) *Picturing Inference*, *Gerechtigkeitswissenschaft*, Berliner Wissenschafts–Verlag, Berlin.

Tillers, P. (2006) *The Dynamic Evidence Page*, Accessed 14–5–2009. http://www.tillers.net/Last.

Toolan, M. (2001) *Narrative: A Critical Linguistic Introduction*, 2nd edition, Routledge, London, New York (New York).

Toulmin, S.E. (1958) *The Uses of Argument*, Cambridge University Press, Cambridge.

Toulmin, S.E. (2003) *The Uses of Argument*, Updated edition, (originally published in 1958), Cambridge University Press, Cambridge.

Trabasso, T. and Sperry, L.L. (1985) Causal relatedness and importance of story events. *Journal of Memory and Language* 24:5, 595–611.

Trabasso, T. and Broek, P. van den. (1985) Causal thinking and the representation of narrative events. *Journal of Memory and Language* 24:5, 612–630.

Twardy, C. (2004) Argument maps improve critical thinking. *Teaching Philosophy* 27:2, 95–116.

Twining, W.L. (2007) Argumentation, stories and generalizations: a comment. *Law, Probability & Risk* 6, 169–185.

Twining, W.L. (1994) *Rethinking Evidence: Exploratory Essays*, Northwestern University Press, Evanston (Illinios).

Twining, W.L. (1995) Anchored narratives, a comment. *European Journal of Crime, Criminal Law and Criminal Justice* 3, 106–114.

Twining, W.L. (1999) Necessary but dangerous? generalizations and narrative in argumentation about "facts" in criminal process. In Nijboer, J.F. and Malsch, M. (eds.), *Complex Cases: Perspectives on the Netherlands Criminal Justice System*, 69–98, Thela Thesis, Amsterdam.

Twining, W.L. (2006) *Rethinking Evidence: Exploratory Essays*, Cambridge University Press, Cambridge.

Van Gelder, T. (2007) The rationale for rationale. *Law, Probability and Risk* 6, 23–42.

Verheij, B. (1996) *Rules, Reasons, Arguments: Formal Studies of Argumentation and Defeat*, Doctoral dissertation, University of Maastricht.

Verheij, B. (1999) Automated argument assistance for lawyers. *Proceedings of the 7th International Conference on Artificial Intelligence and Law*, 43–52, ACM New York (New York).

Verheij, B. (2000) Anchored narratives and dialectical argumentation. In Van Koppen, P.J. and Roos, N. (eds.), *Rationality, Information and Progress in Law and Psychology: Liber Amicorum Hans F. Crombag*, 203–226, Metajuridica Publications, Maastricht.

Verheij, B. (2003a) DefLog: On the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation* 13:3, 319–346.

Verheij, B. (2003b) Dialectical argumentation with argumentation schemes: an approach to legal logic. *Artificial Intelligence and Law* 11:2, 167–195.

Verheij, B. (2005a) Evaluating arguments based on Toulmin's scheme. *Argumentation* 19:3, 347–371.

Verheij, B. (2005b) *Virtual Arguments: On the Design of Argument Assistants for Lawyers and Other Arguers*, T.M.C. Asser Press, Den Haag.

Verheij, B. and Bex, F.J. (2009) Accepting the truth of a story about the facts of a criminal case. In Kaptein, H., Prakken, H., and Verheij, B. (eds.), *Legal Evidence and Proof: Statistics, Stories, Logic*, 161–193, Ashgate, Aldershot.

Verheij, B., Hage, J.C. and Herik, H.J. van den. (1998) An integrated view on rules and principles. *Artificial Intelligence and Law* 6:1, 3–26.

Vreeswijk, G. (1993) Defeasible dialectics: A controversy – oriented approach towards defeasible argumentation. *Journal of Logic and Computation* 3:3, 317–334.

Wagenaar, W.A. and Crombag, H.F.M. (2005) *The Popular Policeman and Other Cases: Psychological Perspectives on Legal Evidence*, Amsterdam University Press, Amsterdam.

Wagenaar, W.A., Koppen, P.J. van and Crombag, H.F.M. (1993) *Anchored Narratives: The Psychology of Criminal Evidence*, St. Martin's Press, New York (New York).

Walton, D.N. (1998) *The New Dialectic: Conversational Contexts of Argument*, University of Toronto Press, Toronto.

Walton, D.N. (2001) Abductive, presumptive and plausible arguments. *Informal Logic* 21:2, 141–172.

Walton, D.N. and Schafer, B. (2006) Arthur, George and the mystery of the missing motive: towards a theory of evidentiary reasoning about motives. *International Commentary on Evidence* 4:2, 1–47.

Walton, D.N. (1996) *Argumentation Schemes for Presumptive Reasoning*, Lawrence Erlbaum Associates, Mahwah (New Jersey).

Walton, D.N. (2002) *Legal Argumentation and Evidence*, Penn, State University Press, University Park (Pennsylvania).

Walton, D.N. and Krabbe, E.C.W. (1995) *Commitment in Dialogue*, State University of New York Press, Albany (New York).

Walton, D.N. and Reed, C.A. (2002) Argumentation schemes and defeasible inferences. *Proceedings of the ECAI 2002 Workshop on Computational Model of Natural Argument*, 45–55.

Walton, D., Reed, C. and Macagno, F. (2008) *Argumentation Schemes*, Cambridge University Press, Cambridge.

Warren, W.H., Nicholas, D.W. and Trabasso, T. (1979) Event chains and inferences in understanding narratives. *New directions in discourse processing* 2, 23–52.

Wigmore, J.H. (1931) *The Principles of Judicial Proof or the Process of Proof as Given by Logic, Psychology, and General Experience, and Illustrated in Judicial Trials*, 2nd edition, Little, Brown and Company, Boston (Massachusetts).

Wigmore, J.H. (1940) *A Treatise on the Anglo – American System of Evidence in Trials at Common Law*, 3rd edition, Little, Brown and Company, Boston (Massachusetts).

Wyner, A., Bench-Capon, T.J.M. and Atkinson, K. (2008) Three senses of argument. In Sartor, G. (eds.), *Computable Models of the Law: Languages, Dialogues, Games, Ontologies*, Springer, Berlin.

# Name Index

# Subject Index

**D**



**E**



**F**



**G**

**S**



**T**



**V**



**W**