When Stories and Numbers Meet in Court

Constructing and Explaining Bayesian Networks for Criminal Cases with Scenarios

Charlotte S. Vlek

The research in this thesis is part of the project "Designing and Understanding Forensic Bayesian Networks with Arguments and Scenarios" in the Forensic Science programme, financed by the Netherlands Organisation for Scientific Research (NWO).

Printed by: Gildeprint

Image credit for cover images:
Fingerprint: https://commons.wikimedia.org/wiki/File:Fingerprint_picture.svg
DNA: https://pixabay.com/en/dna-science-biology-research-718905/
Gun: https://pixabay.com/en/pistol-m9-gun-weapon-silhouette-555656/
Sherlock Holmes: https://pixabay.com/en/sherlock-holmes-detective-147255/

ISBN digital version: 978-90-367-9076-5
ISBN printed version: 978-90-367-9079-6

rijksuniversiteit
groningen

When Stories and Numbers Meet in Court

Constructing and Explaining Bayesian Networks for Criminal Cases with Scenarios

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 28 oktober 2016 om 11.00 uur

door

Charlotte Stephanie Vlek

geboren op 16 juli 1985
te Groningen

Promotores
Prof. dr. H.B. Verheij
Prof. dr. L.C. Verbrugge
Prof. dr. mr. H. Prakken

Copromotor
Dr. S. Renooij

Beoordelingscommissie
Prof. dr. N. Fenton
Prof. dr. P.J. van Koppen
Prof. dr. J.W. Romeijn

Contents

Part I: Introduction and preliminaries

1. Introduction

In a criminal trial, a judge or jury is presented with a collection of evidence. A primary task is one of fact-finding: based on the evidence, a conclusion must be drawn about what really happened. To do this, a judge or jury needs to apply some sort of rational reasoning in order to draw conclusions from the evidence about what may have happened.

With the rise of modern forensic techniques such as DNA profiling, a judge or jury is now faced with the task of reasoning about pieces of evidence that come with some quantified uncertainty. For example, when forensic experts report on a DNA match, they will typically report a so-called *random match probability*: the probability that the match would be found if the suspect were not the source of the DNA trace that was tested. A challenging task for a judge or jury is to take into account the whole case, thereby combining qualitative evidence such as witness testimonies with quantitative evidence.

Throughout this introduction, we use the case of Steven Avery as an illustration of our ideas. The Avery case is a complex case, which recently gained a lot of interest as a result of the Netflix documentary series 'Making a Murderer'.[1] A brief background on this case is as follows: Steven Avery was convicted for a rape that took place in 1985 in Wisconsin, USA. After Avery spent 18 years in prison, another person was identified as the rapist, thanks to advances in DNA profiling techniques. Avery was acquitted in 2003 and filed a civil lawsuit against the local sheriff's department and district attorney, to receive compensation for his wrongful conviction. During the time of the lawsuit, a young female photographer went missing and was last seen photographing a car that was for sale on Avery's property. Avery was charged with the murder and was found guilty. Evidence in this case was complex, but included the remains of the victim's burnt body, found on Avery's property; a match between Avery's DNA profile and some blood found in the victim's car; and the victim's car key found in Avery's bedroom.

The Avery case illustrates the need to interpret quantitative evidence in the context of the other evidence. In the Avery case, the DNA evidence points to Avery's blood being in the victim's car, but this intermediate conclusion requires further interpretation within the case. In particular, Avery always maintained his innocence and claims that he is being framed, possibly by the police as a retaliation for his civil lawsuit. Note that the DNA evidence supports both accounts: it fits the scenario in

[1]Case files can be found on www.wicourts.gov with case number 2005CF000381 (Circuit Court, Manitowoc County) and Appeal No. 2010AP411-CR (the appeal case), or on www.stevenaverycase.org

which Avery is guilty, as well as the scenario in which the blood was planted by the police. To interpret quantitative evidence properly, it is thus crucial that a judge or jury can put the evidence into context.

When reasoning with legal evidence, a judge or jury has various tools to their disposal, which are typically divided into the following three approaches: an argumentative approach, a scenario-based approach and a probabilistic approach (Anderson et al., 2005; Kaptein et al., 2009; Dawid et al., 2011). Each approach will be discussed in more depth in Section 1.1, including advantages and disadvantages of each. In short, argumentation tends to work well with the adversarial nature of reasoning with legal evidence, scenarios put evidence into context with their global perspective on a case and the probabilistic approach provides the formal tools to reason about degrees of uncertainty. A combination of or translation between these various approaches is not straightforward, and is currently a topic of research (Verheij et al., 2016).

In this thesis we propose to combine a probabilistic approach with a scenario-based approach, to enable reasoning with quantitative evidence in the context of a legal case as a whole. This idea will be further motivated in Section 1.2 below. An outline of this thesis can be found in Section 1.3.

1.1 Three approaches to reasoning with legal evidence

As noted above, three main approaches to reasoning with legal evidence exist, with a respective focus on arguments, scenarios or probabilities. In this section, each approach will be discussed briefly, illustrated with examples from the Avery case. The three approaches are then compared and their advantages and disadvantages are discussed.

1.1.1 Arguments

In the argumentative approach to reasoning with legal evidence, arguments can support or attack conclusions based on the evidence. Such an argument typically applies a general rule to some piece of evidence to derive a conclusion. An example from the Avery case can be found in an argument concerning the car key that was found in Avery's bedroom, using the following general rule: 'when an object is found in a person's home, this person usually placed the object there'. The argument is then as follows:

Example (An argument). The victim's car key was found in Avery's bedroom. When an object is found in a person's home, this person usually placed the object there. So Avery placed the victim's car key in his bedroom.

Such rules can be chained to form more complex arguments. For example, the following rule can be applied to the conclusion of the example argument above:

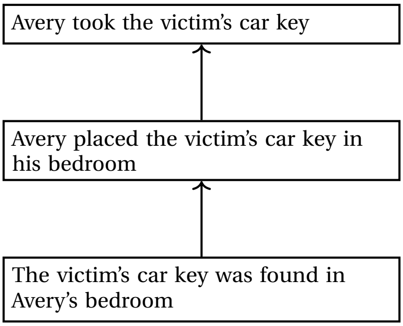


Figure 1.1: An example of an argument.

'when X placed a stolen object in his home, X probably stole this object'. This more complex argument is depicted in Figure 1.1.

An argument can attack another argument, either by disproving the conclusion, or by showing that the rule is not applicable. For example, the aforementioned rule 'when an object is found in a person's home, this person usually placed the object there' does not lead to a correct conclusion when the object was placed there by someone else. There is then an exception to the rule, which can be used to attack the argument. The argument and the attack can be represented as shown in Figure 1.2.

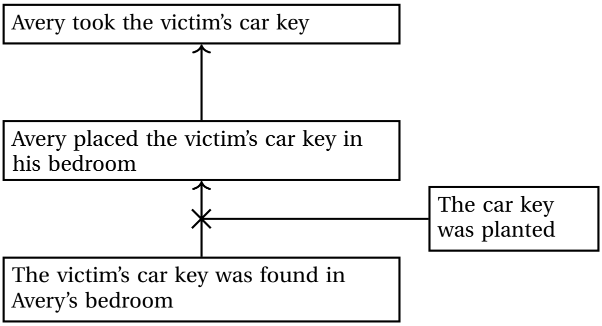


Figure 1.2: An example of attacking arguments.

These ideas on legal argumentation date back to Wigmore (1931), who used diagrams, known as Wigmore charts, in which complex arguments about legal evidence could be analysed. This approach has been further developed by Anderson et al. (2005). Argumentation is furthermore a topic of interest from a computational perspective, leading to the application of techniques from artificial intelligence (AI) to argumentation (Bench-Capon and Dunne, 2007; Van Eemeren et al., 2014) and to argumentation about legal evidence in particular Bex et al. (2003).

The argumentative approach closely fits with the dialectical nature of reasoning with legal evidence, in which the prosecution and the defence form arguments for and against an ultimate claim (Bex, 2011). Characteristic for the argumentative approach is that one should always take into account all arguments for and against a conclusion. The argumentative approach is furthermore intuitive to use for a judge or jury, but it is typically atomistic in nature, focussing on details of the case separately rather than viewing the case as a whole (Bex, 2011).

1.1.2 Scenarios

In the scenario-based approach (also called the narrative approach), several alternative scenarios about what may have happened are formed and compared to find the best scenario. In the Avery case, a scenario for the prosecution can be as follows:

Example (A scenario). The victim was at Avery's property to take pictures of a car that was for sale. She went into Avery's home. Avery then raped her and tried to slit her throat with a knife. He moved her to the garage, where he shot her. He then burned her body in a burn pit behind his house and moved the victim's car to the auto salvage yard on his property.

An alternative scenario for the defence can be:

Example (An alternative scenario). The police wanted to frame Avery because of a pending court case in which Avery wanted compensation for the years he spent in prison. An unknown person killed the victim. The police found the remains of the victim's body and moved it to the Avery property. They moved the victim's car to Avery's salvage yard and placed some of Avery's blood (which they had in a vial as evidence from the previous case) in the car. They placed the key from the victim's car in Avery's bedroom.

These scenarios should now be compared in light of the available evidence. Several researchers pursued the development of a normative framework for reasoning with scenarios. Pennington and Hastie (1993) developed their so-called Story Model, which was not only meant to take into account how to compare scenarios and select 'the best', but also how to relate a scenario to a verdict. Wagenaar et al. (1993) developed the Anchored Narratives Theory, in which a key aspect is the notion that any scenario should be safely anchored in our general knowledge about the world. Recently, Bex (2011) developed a Hybrid Theory of Stories and Arguments, which combines the scenario-based approach with an argumentative approach to reasoning with legal evidence. This resulted in a formal theory in which a scenario can be supported or attacked with arguments. These scenario-based approaches will be discussed in more depth in Chapter 2.

The scenario-based approach is said to help reduce the risk of tunnel vision (Wagenaar et al., 1993), since it requires that multiple alternative scenarios are compared. Forming scenarios furthermore provides a global perspective on a case, taking into account the whole collection of evidence. In this respect, scenarios also serve as a way to make sense of the evidence. This was shown in experiments by Pennington and Hastie (1993), which showed that a judge or jury tends to form scenarios about what happened to place the evidence into context.

On the other hand, when a judge or jury reasons with scenarios, there is a real danger of a good scenario being chosen over a true scenario. Experiments by Bennett and Feldman (1981) revealed that a complete and well-structured scenario was believed more readily by test subjects than an incomplete or unstructured scenario, even if the latter really happened and the former was fake. This led them to conclude that one should be wary of a good scenario pushing out a true scenario, which was later also emphasised by Anderson et al. (2005).

1.1.3 Probabilities

In a probabilistic approach, the probability $\Pr(h)$ of a hypothesis h is of interest, and in particular the probability of h given evidence e, written as $\Pr(h|e)$. Consider the following example in the Avery case:

Example (Posterior probability). Consider the hypothesis h: Avery is the donor of the blood in the victim's car. There is also evidence e: a DNA match was found between the blood in the victim's car and Avery's profile. The probability that is of interest is then $\Pr(h|e)$: the probability that Avery is the donor of the blood in the victim's car given the DNA match.

In this context, the probability $\Pr(h|e)$ is called the *posterior probability* of h after taking evidence e into account. This probability can be calculated using Bayes' rule:

$$\Pr(h|e) = \frac{\Pr(e|h) \cdot \Pr(h)}{\Pr(e)}.$$

In the Avery case, this means that the following probabilities need to be available to calculate $\Pr(h|e)$:

- the *prior probability* $\Pr(h)$: the probability that Avery is the donor of the blood in the victim's car (without taking any evidence into account); and
- the *likelihood* of evidence $\Pr(e|h)$: the probability that the evidence would be found given the assumption that h is the case; and
- the probability $\Pr(e)$ that the evidence would occur. This probability can be calculated from other probabilities, in particular using that $\Pr(e) = \Pr(e|h) \cdot \Pr(h) + \Pr(e|\neg h) \cdot \Pr(\neg h)$ in which $\neg h$ is the complement of h. This requires that $\Pr(e|h)$ as well as $\Pr(e|\neg h)$ are known.

It is said that the posterior probability $\mathrm{Pr}(h|e)$ is found from the prior probability h by *updating* with evidence e. This process is known as Bayesian updating and can be applied repeatedly to update with multiple pieces of evidence. Probabilities will be discussed in more depth in Chapter 2.

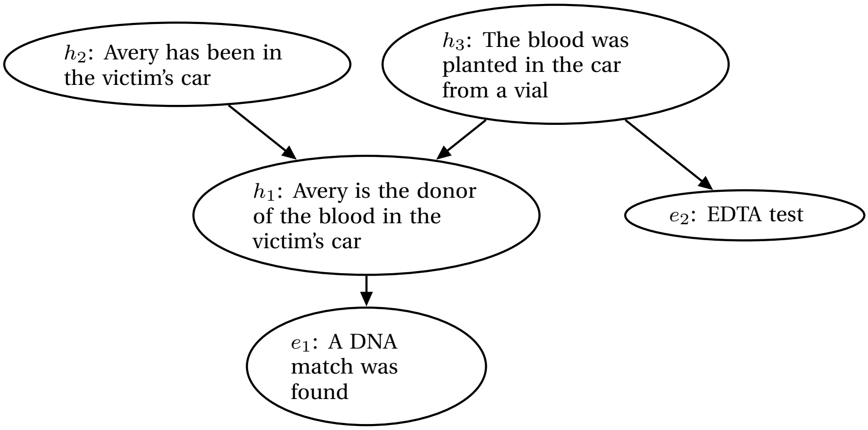


Figure 1.3: A Bayesian network graph for the blood traces in the Avery case.

For a large case, this process of updating can quickly become complex, since relations between the evidence also need to be taken into account. A useful tool to deal with possibly interrelated evidence and hypotheses is a Bayesian network. A Bayesian network consists of a graph and probability tables, together representing a joint probability distribution. The graphical structure contains information about (in)dependencies between variables and from the network any probability of interest can be computed. Bayesian networks will be discussed more elaborately in Chapter 2. An example of the graph of a Bayesian network is shown in Figure 1.3, modelling the following situation in the Avery case:

Example (Interrelated hypotheses and evidence). Consider again the hypothesis and evidence from the previous example: h_1: Avery is the donor of the blood in the victim's car and e_1: a DNA match was found. This can be extended with two further hypotheses: h_2: Avery has been in the victim's car and h_3: the blood was planted in the car from a vial. In an attempt to find evidence for the latter hypothesis, the defence requested a test to find whether the blood in the car contained any EDTA: a chemical that is added to blood in the lab to keep it fluid. This led to evidence e_2: no EDTA was found. The relations between these hypotheses and the evidence are represented in Figure 1.3.

In addition to the graph shown in Figure 1.3, a Bayesian network for this example will have probability tables for each node, such as the probability table for the node

EDTA test, which will indicate the probabilities of a positive/negative test result when the blood was from a vial and the probabilities of a positive/negative test result when the blood was not from a vial.

Bayesian networks have been studied as a tool for probabilistic reasoning about legal cases (Dawid and Evett, 1997; Taroni et al., 2006). A main advantage is that Bayesian networks have a clear calculus for dealing with degrees of uncertainty (Taroni et al., 2006). However, the most common objection against the use of Bayesian networks is that many probabilities need to be specified, which are not always readily available (Jensen and Nielsen, 2007). In some domains a Bayesian network can be learned from data, but in legal applications this is typically not the case. For the assessment of probabilities, elicitation techniques exist to assist an expert in estimating appropriate probabilities based on their expertise (see Renooij (2001) and Chapter 2).

A probabilistic approach is not typically insightful for a judge or jury. This may lead to reasoning errors such as the prosecutor's fallacy, in which the conditional probability of a piece of evidence given a hypothesis is mistakenly thought to be the same as the probability of that hypothesis given the evidence (Thompson and Schumann, 1987). Bayesian networks have been proposed by some as a solution, since the graphical aspect of these networks could be more insightful for a judge or jury (Fenton and Neil, 2011).

In any case, the application of Bayesian methods in court is subject of debate, so more research is required (Fenton, 2011). On the one hand, in the UK the Court of Appeal ruled in 2010 that Bayes' theorem should not be used in evaluating evidence, except for DNA and 'possibly other areas where there is a firm statistical base', which was followed by a lively scientific discussion (see the 2012 special issue of Law, Probability and Risk on the R v T case; Vol. 4, No. 2). On the other hand, in the Netherlands the use of Bayesian thinking has been advocated by an advocate general of the Supreme Court together with the Netherlands Forensic Institute (Nederlands Forensisch Instituut, NFI) (see (Berger and Aben, 2010b,c,a)).

1.1.4 The three approaches compared

The three approaches as described in the previous sections each have their own key characteristics, which are summarised in Table 1.1. Argumentation is close to the adversarial setting of a courtroom, since it uses arguments for and against a certain conclusion. The adversarial setting can be recognised somewhat in the scenario-based approach as well, since this requires that alternative scenarios are compared.

A key characteristic of a scenario-based approach is the global perspective it provides. In this approach, coherent scenarios are used to consider the whole case and compare alternative accounts of what happened.

	Arguments	Scenarios	Probabilities
Adversarial setting	✓	(✓)	-
Global perspective	-	✓	-
Degrees of uncertainty	(✓)	-	✓
Closeness to intuition	✓	✓	-
Standard formalisation	(✓)	-	✓

Table 1.1: Characteristics of the three approaches to reasoning with legal evidence. A check mark ✓ denotes that an approach has this characteristic, while a check mark in parentheses (✓) denotes that the approach has this characteristic only to a limited extent. A hyphen - denotes that the characteristic is absent in this approach.

Typical for the probabilistic approach is that it can work with degrees of uncertainty. The degree of uncertainty can be quantified numerically and these uncertainties can be used to calculate the probability of one specific hypothesis of interest. Some argumentation methods also allow the specification of degrees of uncertainty, in the form of argument strength. However, there is no consensus on how the strengths of various arguments interact when they are combined.

Arguments and scenarios are close to the intuition of a judge or jury, since arguments match well with the adversarial setting and scenarios can be used to make sense of the evidence, by virtue of their global perspective on a case.

The formal development of the three approaches is currently in different stages. In argumentation, various formalisations have been developed, but there is no single commonly accepted framework (Van Eemeren et al., 2014). Scenario-based reasoning is formally less well-developed and only recently received more attention from a computational perspective (Bex, 2011; Shen et al., 2006). The probabilistic approach has a widely accepted formalisation in terms of standard probability calculus (Taroni et al., 2006).

1.2 Problem statement and research questions

As described in the previous section, each of the three approaches to reasoning with legal evidence has certain characteristics. A current topic of research is how a combination or translation between these separate approaches might be established (Verheij et al., 2016). Although the use of probabilities in court is still subject of debate, forensic experts are advocating the use of Bayesian methods (Aitken, 2012; Berger and Aben, 2010b,c,a), which makes it especially relevant to study how a probabilistic approach can be made accessible for a judge or jury. This thesis aims to study a combination of approaches, based on the assumption that such a combination can inherit advantages of each approach.

In previous research, the combination of arguments and scenarios was studied by Bex (2011), which led to the development of his Hybrid Theory of Stories and Arguments. Timmer et al. have worked on the combination of arguments and Bayesian networks (see e.g. (Timmer et al., 2015a)) and Verheij (2014) has studied the combination of all three approaches. An overview of these various research projects and their connections can be found in Verheij et al. (2016). The topic of this thesis is the combination of scenarios and probabilities. This leads to the following problem statement:

Problem statement. How can a scenario-based approach and a probabilistic approach be combined in reasoning with legal evidence in a criminal case, employing advantages of both approaches?

A probabilistic approach is particularly suitable for dealing with the degrees of uncertainty of a case that includes quantitative evidence. It is currently typically only used to analyse parts of a case on a small scale. A scenario-based approach is better suited to consider the case from a global perspective. We aim to combine the two approaches such that the global perspective of a scenario can provide the context for the selection of relevant variables in a probabilistic approach, while reducing the risk of tunnel vision by requiring that several alternative scenarios are taken into account. On the other hand, the probabilistic approach can be beneficial to avoid the common pitfall of selecting a good scenario over a true scenario, by providing a solid formal framework to analyse the evidence for each scenario.

For the purpose of representing a case as a whole, Bayesian networks are potentially a good candidate as a probabilistic tool, since they deal well with interrelated hypotheses and variables in a complex model. Not much work has been done on representing an entire legal case in a Bayesian network, though Kadane and Schum's analysis of the Sacco and Vanzetti case (Kadane and Schum, 1996) presents a probabilistic analysis of a real and complex case, including the use of Bayesian networks. Constructing a Bayesian network for a complex case is not easy. This thesis therefore addresses the following research question:

Research question 1. How can tools from scenario-based reasoning aid the construction of a Bayesian network modelling legal evidence in a criminal case?

In Part II of this thesis, we propose a method which uses scenarios to guide the construction of a network by providing the context and thereby the relevant variables that should go into the network. In particular, we propose the use of ready-made substructures, called *idioms*, which can be used as building blocks to construct a network. The focus of this construction method is on finding the graphical structure of a Bayesian network, while referring to existing elicitation techniques for specifying the probabilities in a network. The construction method results in a network that represents several alternative scenarios and the evidence supporting each scenario.

When using a Bayesian network to model an entire case, it is crucial that a judge or jury gains some understanding of the network. It is especially important that a judge or jury understands which assumptions go into a model, since this is what ultimately determines the outcome (Fenton and Neil, 2011). In this thesis we therefore address as a second research question:

Research question 2. How can tools from scenario-based reasoning aid the understanding of a Bayesian network modelling legal evidence in a criminal case?

In Part III of this thesis, we extend our construction method with techniques to extract from a Bayesian network the alternative scenarios, information concerning their quality and their relations to the evidence. This way, a judge or jury can understand the network in terms of scenario-based reasoning. Simultaneously, the reported information also serves as feedback to check whether the case has been modelled correctly.

The contribution of this thesis is thus a method that combines scenarios and probabilities, by constructing and explaining Bayesian networks with scenarios. Four narrative idioms are provided for representing scenarios and scenario quality in a Bayesian network, as well as a procedure for using scenarios to guide the construction of the network. In this construction, the network is annotated such that scenarios can later be extracted from the network to form a report about the content of the network. This report is enhanced with information about the quality of the scenarios and the relations of the evidence to each scenario. Finally, we test our method in two case studies.

The use of Bayesian networks comes with the issue of having to specify all the required probabilities (Kjærulff and Madsen, 2008). This is a well-known limitation of Bayesian networks and as a result, in our method many probabilities will be based on subjective estimates. Nonetheless, by making these numbers precise, the subjective interpretation of a legal case is made explicit. Our method is thereby intended to formalize a subjective account of a legal case, thus forming a tool for a judge or jury to structure their thoughts rather than a tool for reaching an objective verdict. This research provides foundations for a software tool that can be used by a judge or jury in which they can model their view on a case, with specific probabilities supplied as much as possible by a domain expert.

1.3 Outline of the thesis

This thesis is organised as follows. Chapter 2 in Part I of this thesis serves as an introduction to the topics of Bayesian networks and scenario-based reasoning. In this chapter, basic notions about probabilities, Bayesian networks and reasoning with scenarios are introduced.

Part II discusses the construction of Bayesian networks. In Chapter 3, a method is presented for constructing Bayesian networks for legal evidence with scenarios.

The proposed method includes the use of four newly developed narrative idioms as building blocks for a Bayesian network and an interpretation of the quality of a scenario in a Bayesian network context. Together these provide the foundations needed for the construction method where the narrative idioms and scenario quality are used to gradually construct a Bayesian network for a case.

The construction method is tested on a case study in Chapter 4. This chapter studies the Anjum murders case, which was previously investigated by Crombag and Israëls (2008) and also served as a case study for Bex (2011).

Part III is about the understanding of Bayesian networks. In Chapter 5, the method from Part II is extended such that after construction, scenarios as they are modelled in the network can be extracted. It also includes a discussion on how to report on scenario quality and evidential support as represented in the network. Finally, a reporting format is provided in which each scenario together with its scenario quality and evidential support can be reported to a judge or jury.

The extended method is tested on a case study in Chapter 6. The case in this chapter is one in which statistical reasoning clearly played a role in the decision. In this case, a suspect is thought to have helped move the body of a murder victim. An important part of the appeal case was concerned with how likely the evidence is given several alternative hypotheses. The appeal of this case can be found (in Dutch) on www.rechtspraak.nl using code ECLI:NL:GHARL:2014:8941.

Part IV concludes this thesis with a chapter on related work (Chapter 7) and a conclusion (Chapter 8).

1.4 Provenance of chapters

This thesis is based on several published papers. A list of all publications can be found on page 173. In this section we indicate for the main body, Chapters 3-7 in Parts II and III of this thesis, on which paper(s) each chapter is based. The chapters in Part I (Introduction and Preliminaries) and Part IV (Related work, discussion and conclusion) are more general and hence based on all research that was performed during the project.

The construction method (Chapter 3) is mainly based on Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014a). Building Bayesian networks for legal evidence with narratives: A case study evaluation. *Artificial Intelligence and Law*, 22(4):375 – 421. The chapter also uses elements of Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law.* In press. These journal papers are in turn based on preliminary work that was published in various conference papers, see page 173 for a full list.

The case study in Chapter 4 is based on the case study in Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014a). Building Bayesian networks for legal evidence with narratives: A case study evaluation. *Artificial Intelligence and Law*, 22(4):375 – 421, but is now updated to reflect the latest developments of our method.

The extension of the method with explanation techniques (Chapter 5) was first proposed as a separate explanation method in Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law*. In press. That paper is also the original source of the case study in Chapter 6.

2. Background: Bayesian networks and scenarios

This chapter provides an introduction to concepts that are used throughout the thesis. Section 2.1 introduces basic notions from probability theory, Section 2.2 provides background on Bayesian networks and Section 2.3 provides a description of scenario-based reasoning.

2.1 Probabilities

When working with probabilities in the area of evidential reasoning, we are typically interested in the probability of a certain hypothesis (such as ‘the suspect committed the crime’), taking into account the evidence we have (Taroni et al., 2006). In this section we introduce basic notions and notations needed to discuss these probabilities.

Probabilities The probability of a hypothesis h is denoted as $\Pr(h)$ and is given by a number between 0 and 1 (possibly equal to 0 or 1).

Example (Probability of a hypothesis). Consider a criminal case which includes the following hypothesis h: the suspect is the donor of the blood traces found at the crime scene. The probability of this hypothesis is given by $\Pr(h)$.

In what follows, we work with stochastic variables. A stochastic variable is a variable of which the value is unknown, but which can take on different values that are each assigned some probability, together adding up to 1.

Example (Variables). Consider a case in which it needs to be established whether the suspect is the donor of the blood traces found at the crime scene. We introduce a variable H, which can take two values: h_1: the suspect is the donor of the blood traces at the crime scene, and h_2: the suspect is *not* the donor of the blood traces at the crime scene. The probability that the suspect is the donor of the blood traces at the crime scene is given by the probability that the variable H takes value h_1. This is denoted by $\Pr(H = h_1)$, or in shorthand notation $\Pr(h_1)$. Similarly, the probability that the suspect is not the donor is given by $\Pr(H = h_2)$ or $\Pr(h_2)$.

We write $\Pr(H)$ for the probability distribution of variable H, which specifies the probability $\Pr(H = h_i)$ for all the values h_i of H.

For a stochastic variable H, its values $h_1, ...h_n$ must be mutually exclusive and exhaustive. This means that the variable can only have one value at a time (mutually exclusive) and the variable must take one of these values (exhaustive). For the above example, the outcomes are indeed mutually exclusive and exhaustive when viewed from the perspective of classical logic: the suspect cannot be the donor *and* not the donor simultaneously (the law of non-contradiction), and the suspect must either be the donor or not (the law of the excluded middle) (Barwise and Etchemendy, 1999).

Conditional probabilities In this thesis we will often be interested in a probability of a hypothesis $H = h$ (or in short: h) given the information about some evidence $E = e$ (or in short: e). This can be expressed as a conditional probability, namely the probability of h, conditioned on e: $\Pr(h|e)$. For two variables A and B, one can look at the conditional probability of a conditioned on b:

$$\Pr(a|b).$$

This number is in the collection of conditional probability distributions $\Pr(A|B)$, which specifies the conditional probability for all configurations of values that the variables A and B might take.

Example (Conditional probability). Suppose a DNA match is found, which shows that the suspect's profile matches the blood found at the crime scene. Let us denote the match as e: the suspect's profile matches the blood traces found at the crime scene. A probability of interest is now the conditional probability that the suspect is the donor, given this match: $\Pr(h|e)$.

Joint probabilities When there are two (or more) variables A and B, the probability of values a and b of these variables occurring together is denoted by the joint probability $\Pr(a, b)$.

Example (Joint probability). Suppose that at the crime scene a bullet is found (evidence e_1), as well as a bloody knife (evidence e_2). The joint probability $\Pr(e_1, e_2)$ describes the probability that these two pieces of evidence would occur together.

The joint probability distribution (JPD) $\Pr(A, B)$ specifies the joint probability for any configuration of values of the variables.

The joint probability relates to the conditional probability as follows:

$$\Pr(a, b) = \Pr(a|b) \cdot \Pr(b).$$

Conditional independence When the probability distribution of variable A does not change when it is conditioned on some other variable B, these variables are said to be independent. That is, when $\Pr(a_i|b_i) = \Pr(a_i)$ for all values a_i of A and all values b_i of B, then A and B are said to be conditionally independent, written as $\Pr(A|B) = \Pr(A)$. Alternatively, A and B are independent when $\Pr(A, B) = \Pr(A) \cdot \Pr(B)$.

When variables A and B are independent when conditioned on a third variable C, they are said to be conditionally independent given C. This is the case when

$$\Pr(a_i|c_i, b_i) = \Pr(a_i|c_i)$$

for all values a_i of A, b_i of B and c_i of C. This is written as $\Pr(A|C, B) = \Pr(A|C)$.

Prior and posterior probabilities Using the notion of conditional probabilities, one can think of the conditional probability $\Pr(h|e)$ as an updated belief for the hypothesis h, updated with the knowledge about e. In this context, the probability of the hypothesis before updating is $\Pr(h)$ and is called the prior probability of h. The conditional probability $\Pr(h|e)$ is called the posterior probability of h after updating with evidence e.

Example (Prior and posterior probabilities). Consider the example case with h: the suspect is the donor of the blood traces at the crime scene and e: the suspect's profile matches the blood traces found at the crime scene. $\Pr(h)$ is the prior probability that the suspect is the donor of the blood traces. Updating the prior probability with evidence e leads to the posterior probability $\Pr(h|e)$, which is the probability that the suspect is the donor given the DNA match that was found.

Bayes' rule Bayes' rule can be used to calculate the posterior probability given the prior probability, as follows:

$$\Pr(h|e) = \frac{\Pr(e|h) \cdot \Pr(h)}{\Pr(e)}.$$

There, the probability $\Pr(e|h)$ is called the likelihood of the evidence. When comparing two hypotheses h_1 and h_2, the odds-version of Bayes' rule can be useful:

$$\frac{\Pr(h_1|e)}{\Pr(h_2|e)} = \frac{\Pr(e|h_1)}{\Pr(e|h_2)} \cdot \frac{\Pr(h_1)}{\Pr(h_2)}.$$

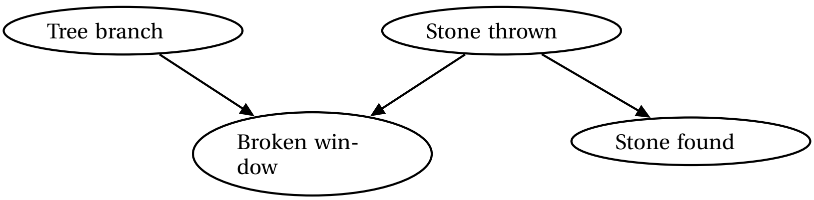
The term $\frac{\Pr(h_1|e)}{\Pr(h_2|e)}$ is now called the posterior odds and the term $\frac{\Pr(h_1)}{\Pr(h_2)}$ denotes the prior odds. The term $\frac{\Pr(e|h_1)}{\Pr(e|h_2)}$ is called the likelihood ratio and is used to find the posterior odds from the prior odds.

2.2 Bayesian networks

A Bayesian network is a compact way of representing a joint probability distribution over a set of discrete stochastic variables. It consists of a directed acyclic graph with nodes representing the variables and conditional probability tables for each node in the graph (Jensen and Nielsen, 2007). An example of a Bayesian network is shown in Figure 2.1. This example network models four variables: a window that is either broken or not (variable `Broken window`), a tree branch that may or may not have hit the window due to heavy wind (variable `Tree branch`), a stone that may or may not have been thrown through the window by someone (variable `Stone thrown`) and a stone that may or may not be found inside the house (variable `Stone found`). Each of these variables has possible values true (`T`) and false (`F`).

Tree branch	
`Tree branch = T`	0.001
`Tree branch = F`	0.999

Stone thrown	
`Stone thrown = T`	0.01
`Stone thrown = F`	0.99



Broken window				
`Tree branch`	`T`		`F`	
`Stone thrown`	`T`	`F`	`T`	`F`
`Broken window = T`	0.99	0.1	0.9	0.001
`Broken window = F`	0.01	0.9	0.1	0.999

Stone found		
`Stone thrown`	`T`	`F`
`Stone found = T`	0.9	0.001
`Stone found = F`	0.1	0.999

Figure 2.1: An example of a Bayesian network.

Nodes Each node in the graph of a Bayesian network represents a stochastic variable, with a set of possible values $\{a_1, ..., a_n\}$, which are mutually exclusive and exhaustive. In the example network from Figure 2.1, all variables have only two possible values, namely true (`T`) and false (`F`). Throughout this thesis we will only use such boolean variables with values true and false. As a result, variables are often formulated as hypotheses (such as 'Stone found') that relate to the variable being assigned value true.

Arrows The arrows in the graph of a Bayesian network specify how variables are related. From these arrows, (in)dependencies between variables can be read (see below for more on on this and the notion of d-separation). When there is an arrow from variable A to variable B, then A is said to be the parent of B and B is the child of A. Child nodes of B are said to be descendants of A, as are children of the children of these nodes and so on. Similarly, parents of A and parents of parents of A (and so on) are ancestors of B.

Probabilities Each node X in a Bayesian network has a conditional probability table (CPT), which specifies conditional probability distributions for X given the different combinations of values for its parents in the network. The conditional probability tables in Figure 2.1 can be read as follows, illustrated with the table for the node `Broken window`. On the left, the table lists the various values of the variable `Broken window`. To read the probability of `Broken window` taking value true (`T`) given a certain configuration of parents, one should find the column which specifies this configuration of parent nodes. For instance, the configuration of `Tree branch = T` and `Stone thrown = T` is the first column that contains numbers. In this case, the probability of `Broken window = T` given `Tree branch = T` and `Stone thrown = T` is 0.99.

When a node has no parents, the probability table specifies the node's prior probability distribution. For example, in Figure 2.1, the probability table for `Tree branch` specifies that $\Pr(\texttt{Tree branch = T}) = 0.001$, while the probability table for `Broken window` specifies that $\Pr(\texttt{Broken window = T}|\ \texttt{Tree branch = F}, \texttt{Stone thrown = T}) = 0.9$.

From the probability tables of a Bayesian network, any probability of interest can be calculated. For instance, the marginal probability distribution of the node `Broken window`, which specifies the probability of a broken window being true or false in general without conditioning on other variables, can be calculated based on the probability tables of `Tree branch`, `Stone thrown` and `Broken window`. Moreover, based on the network, a full joint probability distribution (JPD) over all variables in the network can be retrieved. In fact, a Bayesian network is a compact representation of a JPD since, rather than specifying all joint probabilities of combinations of variables and their values, a Bayesian network only requires the conditional probability of a node conditioned on its parents. Together with the structural information in the graph, this is sufficient to compute a JPD $\Pr(U)$, in which $U = \{A_1, ...A_n\}$ is the set of variables. This is done using the chain rule for Bayesian networks (Jensen and Nielsen, 2007), which can be derived from the general chain rule from probability theory in combination with the independence properties of a Bayesian network ($\text{pa}(A_i)$ denotes the set of parents of A_i):

$$\Pr(U) = \Pi_{i=1}^{n} \Pr(A_i|\text{pa}(A_i)).$$

For the example network in Figure 2.1, this means that

$$\Pr(\texttt{Tree branch, Broken window, Stone thrown, Stone found}) =$$
$$\Pr(\texttt{Broken window}|\texttt{Tree branch}, \texttt{Stone thrown}) \cdot \Pr(\texttt{Stone found}|\texttt{Stone thrown})$$
$$\cdot \Pr(\texttt{Tree branch}) \cdot \Pr(\texttt{Stone thrown}).$$

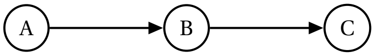
Instantiating nodes A Bayesian network can be used to calculate the probability of a certain hypothesis given a set of evidence. For example, in the network from Figure 2.1 one may want to calculate the probability that a stone was thrown given the evidence that a window is broken: $\Pr(\texttt{Stone thrown = T}|\texttt{Broken window = T})$. To this end, the information that `Broken window = T` is entered in the network and it is said that the value has been observed and the variable `Broken window` is now instantiated.

To perform calculations within the network, Bayesian network software is available, such as GeNIe[1], SamIam[2] or AgenaRisk[3].

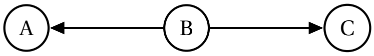
d-Separation From the graph, information can be read about which variables possibly have an influence on each other. This influence may change due to observations of certain variables. When the value of a variable is known, the variable is instantiated in the network. In the network from Figure 2.1 an example of changing influences as a result of instantiated variables is the following: when no observations have been made yet, the finding that there is a broken window (`Broken window = T`) will influence our belief that a stone may be found inside the house (`Stone found = T`). But when we already observed that someone threw a stone at the window (`Stone thrown = T`), the finding that the window is broken (`Broken window = T`) no longer matters for our belief that a stone might be found inside (`Stone found = T`). To denote whether there is possibly an influence between two variables given the set of evidence, the notion of d-separation is used. Whether two variables A and C are d-separated, depends on the chains of nodes between them. A chain between A and C is a sequence of nodes such that for each pair of consecutive nodes in the chain, there is an arrow between these nodes. Note that the direction of the arrow between consecutive nodes is not restricted. Three possible types of connections that may occur within the chain between A and C are shown in Figure 2.2: a serial connection, a diverging connection and a converging connection. Depending on these connections, the chain is either blocked or active (Kjærulff and Madsen, 2008):

Definition 2.2.1 (blocked chain / active chain)**.** A chain between nodes X and Y is *blocked* given a set of instantiated nodes if one of the following holds:

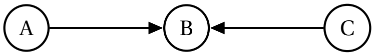
[1]Available on www.bayesfusion.com
[2]Available on reasoning.cs.ucla.edu/samiam/
[3]Available on www.agenarisk.com



(a) A serial connection



(b) A diverging connection



(c) A converging connection

Figure 2.2: Possible connections within a Bayesian network.

- the chain contains a serial connection of three consecutive nodes A, B and C and B is instantiated; or
- the chain contains a diverging connection of three consecutive nodes A, B and C and B is instantiated; or
- the chain contains a converging connection of three consecutive nodes A, B and C and neither B nor any of B's descendants are instantiated.

A chain is *active* given a set of instantiated nodes when it is not blocked given these instantiations.

According to this definition, a serial connection (Figure 2.2(a)), forms an active chain as long as the intermediate node B is uninstantiated, in which case there can be an influence between A and C. The chain becomes blocked as soon as B is instantiated. Similarly, a diverging connection (Figure 2.2(b)) forms an active chain as long as B is uninstantiated. This is the situation as described for the example network in Figure 2.1: as long as we do not know that the stone was thrown, information about a broken window influences our belief that a stone may be found inside the house. As soon as we know for sure whether a stone was thrown (`Stone thrown` is instantiated), the influence between the broken window (`Broken window`) and the stone inside the house (`Stone found`) stops. Finally, a converging connection (shown in Figure 2.2(c)) forms a special case: as long as B and all its descendants are uninstantiated, the chain is blocked, and there is no influence between A and C. But as soon as any information about B becomes available, either by instantiating B directly or by instantiating any of its descendants, the chain between A and C

becomes active. An example of this can be found in Figure 2.1 as well, with A: `Tree branch`, B: `Window broken` and C: `Stone thrown`. As long as we have no information about a broken window, there is no influence between a stone being thrown and a tree branch hitting the window. However, as soon as we have an indication that the window is broken, the tree branch and the stone being thrown will act as competing causes for the broken window. When our belief in a stone being thrown increases (for example, because a stone was found inside the house), our belief in a tree branch hitting the window will decrease. This phenomenon is called intercausal inference or explaining away (Kjærulff and Madsen, 2008).

With the concept of active and blocked chains we can now define d-separation (Kjærulff and Madsen, 2008):

Definition 2.2.2 (d-separation)**.** Two variables X and Y are *d-separated* given a set of instantiated nodes when all chains between X and Y are blocked given these instantiations. When there is at least one active chain between X and Y, the nodes are *d-connected.*

From the structure of the graph we can thus read whether variables X and Y are d-separated. When X and Y are d-separated, a change in belief concerning the value of variable X has no effect on our belief in the value of Y, and X and Y are probabilistically independent. Moreover, when two variables are d-separated given the instantiated node Z, it follows that X and Y are conditionally independent given Z. The converse does not hold: when variables are conditionally independent, they are not necessarily d-separated in the network. The graph of a Bayesian network is therefore a representation of the independence relations amongst its variables, but not necessarily a perfect representation.

Constructing a Bayesian network graph The graph of a Bayesian network is often constructed such that the arrows represent causal relations. However, there is no requirement that an arrow in a Bayesian network should represent causality (Dawid, 2009); the only requirement on a network is that independences are represented correctly. Finding a correct representation is not an easy task, but a causal approach is a useful heuristic in that respect (Jensen and Nielsen, 2007). In our example network in Figure 2.1, the arrows were directed causally.

To further simplify the construction of a Bayesian network graph, methods have been developed (see, e.g., Hepler et al. (2004); Fenton et al. (2013)). These methods mainly focus on the use of recurrent patterns in various networks. Such recurrent patterns are connected to an object-oriented approach to Bayesian networks (Koller and Pfeffer, 1997), in which a network is built from objects which can either consist of single nodes, or of more complex structures of other objects. By defining objects for often recurring substructures, the construction process is simplified.

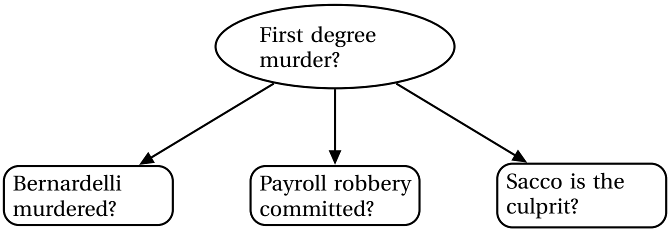


Figure 2.3: A top-level object-oriented network structure for the murder charge against Sacco from Hepler et al. (2004).

Hepler et al. (2004) further studied the specific application of object-oriented Bayesian networks to legal cases. They proposed to combine an object-oriented approach with the use of Wigmore charts from argumentation theory (Wigmore, 1931), borrowing from it the process of dividing the so-called ultimate probandum (the ultimate hypothesis about whether the suspect committed the crime) into penultimate probanda (questions that need to be answered to find out whether the suspect committed the crime). Using the well-known Sacco and Vanzetti case (Kadane and Schum, 1996), Hepler et al. illustrate their approach with the following example: to find whether Sacco is guilty of first-degree murder in the slaying of Alessando Berardelli during a robbery, one must find whether Berardelli died of gunshot wounds, whether he was in possession of payroll at the time he was shot, and whether it was Sacco who intentionally fired shots that took Berardelli's life during the robbery of payroll. With an object-oriented approach a Bayesian network can now be constructed top-down as shown in Figure 2.3, with the ultimate probandum as a single root node and each penultimate probandum as a module (depicted as a rectangular node) that requires further specification.

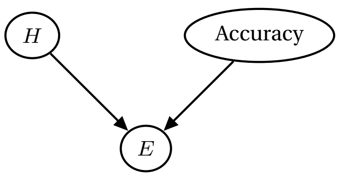


Figure 2.4: The evidence accuracy idiom from Fenton et al. (2013)

To further assist a modeller, Hepler et al. also specified a few often recurring substructures that may occur within the network, such as contradiction or corroboration between evidence and explaining away. Building on this, Fenton et al. (2013) proposed a list of recurrent substructures, called legal idioms. Their evidence accuracy idiom is shown in Figure 2.4, the other idioms can be found in Figure 7.9

in Chapter 7. The list of legal idioms includes idioms for modelling evidence related to a hypothesis, the accuracy of evidence and dependency between evidence, as well as situations more specific to the legal field such as motive, opportunity and alibi evidence.

Eliciting the probabilities in a Bayesian network When constructing a Bayesian network model, there are typically many probabilities that need to be specified. Unless a network can be learned from an available collection of data, the elicitation of probabilities is often not an easy task, since such numbers may not (all) be readily available. Various techniques exist for the elicitation of probabilities (see Renooij (2001)). Many of these techniques are aimed at helping experts in specifying probabilities. This builds on the idea that a probability regarding a certain domain (for example, shoe marks left at a crime scene) needs to be specified by an expert in that specific domain (for example, a forensic analyst with expertise in shoe marks). However, it is unclear how well these techniques work for eliciting probabilities regarding common sense knowledge, since this is not a domain that has clear experts.

Renooij (2001) argues that a domain expert should always receive extensive guidance in the process of eliciting probabilities. For example, questions can be structured to help an expert to compare the probabilities in question to known situations, such as a lottery or a probability wheel (do you think the event is equally likely as a pointer next to a spinning wheel ending up in an area that is one tenth of the area of the wheel?). Additionally, multiple experts can be asked for the elicitation of a single probability, such that their estimations can be compared and possibly averaged.

Finally, a sensitivity analysis can be a useful tool to establish what effect different estimations of a probability may have on the final outcome of the network (Kjærulff and Madsen, 2008). A sensitivity analysis is always performed relative to a hypothesis of interest (for example, the suspect committed the crime) and a set of evidence. Such an analysis sometimes simplifies the elicitation of numbers, since it can be shown that some probabilities require less precise estimations, since their impact on the hypothesis of interest is low. For example, when some probability in the network is estimated to be somewhere in the range of 0.6-0.8, a sensitivity analysis can be used to find how different numbers in this range impact the probability of the hypothesis. When the impact is low, it can be concluded that any number in this range is sufficiently precise, relative to the hypothesis of interest given the evidence.

2.3 Reasoning with scenarios

In a scenario-based approach (or narrative approach) to reasoning with legal evidence, scenarios describe what may have happened concerning a supposed crime (Bennett and Feldman, 1981; Pennington and Hastie, 1993; Wagenaar et al., 1993; Bex, 2011). Several alternative scenarios are compared, thereby considering several alternative accounts of what may have happened. Finally, the 'best' scenario is selected. Such scenario-based approaches are often a form of inference to the best explanation (IBE), in which various alternative explanations of a set of evidence are compared and the best explanation is selected (Pardo and Allen, 2008). Scenario-based approaches are also closely related to coherence approaches to reasoning, such as the theory of explanatory coherence by Thagard (1989) in which explanations of evidence are comprised of several hypotheses and undirected links between these hypotheses express how well these hypotheses cohere.

2.3.1 Scenarios

Scenarios and subscenarios A scenario is a coherent collection of states, events or subscenarios (following Bex (2011)). The basic elements of a scenario are propositions describing states and events, possibly organized in substructures, called subscenarios. Such an hierarchical ordering in subscenarios can also be found in Wagenaar et al. (1993). For example, consider the following scenario:

Example (A scenario). 'John and Mike had a fight. John grabbed a knife. John stabbed the knife into Mike's stomach. John ran off.'

The events 'John grabbed a knife' and 'John stabbed the knife into Mike's stomach' together form a subscenario that describes the events of John stabbing Mike.

Coherence Not just any collection of propositions describing states and events makes a good scenario. A good scenario should be coherent. Consider the following example:

Example (An incoherent collection of events). 'John broke a window. John was a teacher. John lived in a holiday home.'

This is clearly not a coherent scenario: the separate propositions seem unrelated. The example about John and Mike in the previous example also consists of separate propositions, but there they form a coherent whole: a scenario about a fight resulting in a stabbing.

There is no clear consensus on what makes a scenario coherent. In the work by Bennett and Feldman (1981), a good scenario has a clearly identifiable central action and a setting from which the central action seems plausible. Similarly, in their Anchored Narratives Theory, Wagenaar et al. (1993) state that a good scenario contains no contradictions, it has a central action and other elements from the scenario unambiguously explain how and why the central action took place. Finally, Pennington and Hastie (1993) describe in their Story Model that a coherent scenario is complete, consistent and plausible. These notions were later adopted and formalized by Bex (2011) in his Hybrid Theory of Stories and Arguments.

Consistency A scenario is consistent in Pennington and Hastie (1993) when it contains no internal contradictions. Note that this is a factor that is also mentioned by Wagenaar et al. (1993) and later formalized by Bex (2011).

Plausibility The plausibility of a scenario in Pennington and Hastie (1993) is the extent to which it is in line with the decision maker's knowledge about the world. In the Anchored Narratives Theory (Wagenaar et al., 1993), the idea that a scenario should relate to our knowledge of the world is made explicit in the anchoring of scenarios. In the Anchored Narratives Theory, elements of a scenario can be anchored by supplying a general rule that underlies them, such as 'police officers generally do not lie'. Only a scenario that can be sufficiently anchored in so-called 'safe anchors' should be taken into consideration. Inspired by these ideas, Bex formalized plausibility using argumentation. In his Hybrid Theory, elements and generalisations in a scenario can be supported by arguments. An example of this is shown in Figure 2.5, in which evidence of a neighbour overhearing John and Mike fight supports the first element of the scenario, while the common sense knowledge that perpetrators usually run away after a stabbing supports the connection between 'John stabbed Mike' and 'John ran off'. In Bex's formalisation, plausibility is the extent to which a scenario is supported or attacked by arguments based on our knowledge of the world.

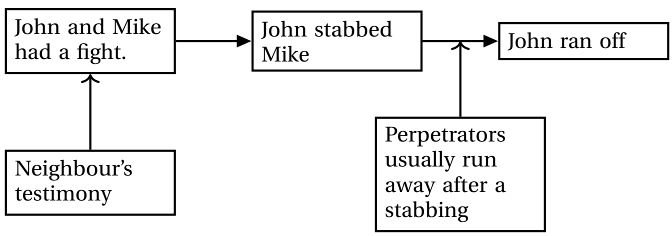


Figure 2.5: An example of a scenario in which elements and generalisations are supported by arguments as in the Hybrid Theory from Bex (2011). The closed arrowheads depict relations within the scenario, while the open arrowheads depict supporting arguments.

Scenario scheme The notion of completeness builds on the idea that a scenario has some underlying structure, such as a story grammar (Rumelhart, 1975), script

(Schank and Abelson, 1977) or scheme (Bex, 2011). For Pennington and Hastie, an example of such an underlying structure is that of an intentional action, which consists of initiating states and events, goals, actions and consequences. A scenario scheme can also be more context specific, such as for example a scenario about a typical stabbing: X was in a fight with Y, X stabbed Y, X ran off.

Completeness Pennington and Hastie (1993) speak of completeness as a scenario having 'all its parts'. This was later formalised by Bex (2011), such that a scenario can be complete with respect to some specific scenario scheme. A scenario is then said to be complete when it fits a scheme (for all elements of the scenario, there is some corresponding element in the scheme) and completes a scheme (for all elements of the scheme, there is some corresponding element in the scenario). The example scenario about John stabbing Mike fits and completes the scheme when the events 'John grabbed a knife' and 'John stabbed the knife into Mike's stomach' are considered as a subscenario that matches the element 'X stabbed Y' in the scheme.

Summary In what follows, we will use shared ideas by Pennington and Hastie (1993), Wagenaar et al. (1993) and Bex (2011), for which the terms completeness, consistency and plausibility can be summarised as follows:

- **Completeness**: A scenario is complete when it 'has all its parts'. This can be formalized with scenario schemes as done by Bex (2011). A scenario is then complete when there is some scenario scheme such that the scenario fits and completes the scheme.
- **Consistency**: A scenario is consistent when it contains no internal contradictions.
- **Plausibility**: The plausibility of a scenario is the extent to which it matches our common sense knowledge about the world. Bex (2011) speaks of the plausibility of elements and connections in a scenario, which is formalized using arguments. The plausibility of an element or connection is then the extent to which it is supported or attacked by arguments based on common sense knowledge.

2.3.2 Comparing scenarios

Comparing alternative scenarios leads to the selection of a 'best' scenario and thus a conclusion about, for example, the suspect's guilt. The criteria by which scenarios should be compared, can roughly be divided into two main factors: the quality of the scenario itself and the relations between the scenario and the evidence. These factors can be found (in one form or another) in Wagenaar et al. (1993), Pennington and Hastie (1993) and Bex (2011).

Quality of a scenario To compare the quality of scenarios, we use the notion of coherence as described in Section 2.3.1. The coherence of a scenario then depends on three factors: completeness, consistency and plausibility. While completeness and consistency are boolean notions, plausibility can be interpreted as being more gradual.

Specifically, a scenario is either complete or it is incomplete and when it is incomplete it is not considered as a real alternative. Similarly, when a scenario contains an inconsistency, it is not considered a good alternative. However, a scenario can be plausible to a certain degree and even a very implausible scenario can still be a good alternative. This is because the plausibility of a scenario interacts with the evidence provided for a scenario: even when a scenario is very unbelievable at first, it may become believable when sufficient evidence is provided (Wagenaar et al., 1993).

Evidential support Pennington and Hastie (1993) consider scenarios to be explanations of the evidence and thus consider the evidence that is ‘covered’ by the scenario, speaking of a scenario’s evidential coverage. In a similar fashion, van Koppen (2011), in his work building on the Anchored Narratives Theory, proposes to compare how well each scenario predicts the evidence. Bex (2011) speaks of the evidential support of the evidence for the scenario. There, arguments based on the evidence can support elements of the scenario. In each of these approaches, there is no clear way in which the evidential support (or coverage) of one scenario can be compared to the evidential support (or coverage) of another. Clearly, if one scenario is supported by evidence E_1 and E_2, while another scenario is supported by E_1, E_2 and E_3, the latter has better evidential support. But if the evidence supporting one scenario is not a superset of the evidence supporting another scenario, there is no clear way to decide which has the best evidential support. After all, one scenario might be supported by a strong DNA match, while the other scenario might be supported by a witness testimony of a very unreliable witness. As we will see later in this thesis, the narrative notion of evidential support can be defined in terms of probabilities, making it possible to compare the amount of evidential support for each scenario.

A good scenario versus a true scenario In experiments by Bennett and Feldman (1981), it was found that participants tend to believe more in a scenario with a coherent structure (a good scenario) than in a scenario that is unstructured or ambiguous in its structure, even when the latter is the one that really took place (the true scenario). This shows that when comparing alternative scenarios, jurors might tend to focus on the quality of a scenario more than on the evidential support for it. This phenomenon was later described as a good scenario pushing out the true scenario by Anderson et al. (2005).

Trade-off In the end, the choice of one scenario over an alternative scenario is a trade-off between scenario quality and evidential support. Consider the example of the Rijkbloem case from Crombag et al. (1992):

Example (The Rijkbloem case). Rijkbloem had a fight with his ex-girlfriend and her parents (the Lammerts family) and in the heat of the moment, the father of the ex-girlfriend (Mr. Lammerts) got shot. Two scenarios need to be compared: Mrs. Lammerts and her daughter (the ex-girlfriend) tell the police that Rijkbloem simply pulled a gun and shot Mr. Lammerts, while Rijkbloem tells the police that Mrs. Lammerts pulled a gun and it accidentally went off in a struggle, the bullet hitting Mr. Lammerts.

Clearly, the scenario about Rijkbloem pulling a gun sounds much more believable. However, with sufficient evidence, the other scenario can be believable as well. In this case, the bullet hit Mr. Lammerts in the nose at a very strange angle, which supports the second scenario (Mrs. Lammerts held the gun) more than the first scenario (Rijkbloem deliberately shot). There is thus one scenario of good quality but with little evidential support, and one scenario of lower quality with more evidential support. It is now up to the judge or jury to decide between these scenarios.

2.3.3 Characteristics of the scenario-based approach

In experiments, Pennington and Hastie (1993) showed that jurors tend to form scenarios about what happened to make sense of the evidence. This way, a scenario provides a global perspective on a case. Several other properties of scenarios are of interest to an application in reasoning with legal evidence and will be discussed below.

Preventing tunnel vision A key element of the scenario-based approach is the comparison of several alternative scenarios. By deliberately considering multiple alternatives, the risk of tunnel vision among jurors is reduced (Wagenaar et al., 1993). Specifically, next to a 'guilty scenario' in which the suspect committed the crime, a judge or jury should try to take an 'innocent scenario' into consideration, in which the suspect is innocent (van Koppen, 2011). This way, the real possibility that the suspect is innocent is taken into account, hopefully preventing the pitfall of focussing too much on the guilty scenario only.

Unfolding scenarios Although a scenario gives an account of what may have happened, there are multiple ways to describe the same events, depending on the level of detail that is chosen to describe what happened. In a legal setting, this can be used to elaborate on details that matter for the particular case and skim over other details that are irrelevant. For example, consider once again the following scenario:

Example (A scenario with details about the stabbing). 'John and Mike had a fight. John grabbed a knife. John stabbed the knife into Mike's stomach. John ran off.'

This scenario specifies some detail of the stabbing (John grabbed a knife, then stabbed Mike in the stomach), while it does not specify any details about the fight between Mike and John. An alternative scenario about the same events could have been the following:

Example (A scenario with details about the fight). 'Mike took John's car without permission. John got angry with Mike. John stabbed Mike. Mike ran off.'

The latter account of what happened includes a more detailed subscenario concerning the fight, while the subscenario about the stabbing is now reduced to a single proposition. Using subscenarios, parts of a scenario can be elaborated to provide more detail when required for the specific case. When elaborating on the details of a certain element, this is called the unfolding of that element.

Transfer of evidential support Due to a scenario's coherence, the events in a scenario form a cluster of elements that somehow belong together. A clear effect of this can be observed: as soon as evidence is available for an element of a scenario, the other elements of a scenario become more believable as well. Consider once again the example scenario:

Example (A scenario). 'John and Mike had a fight. John grabbed a knife. John stabbed the knife into Mike's stomach. John ran off.'

If evidence becomes available that increases our belief in John stabbing Mike with a knife, we will also have an increased belief in John running off afterwards, because we believe this to be a typical collection of events as they may occur together. We will call this phenomenon *transfer of evidential support*, since the support provided by the evidence for John stabbing Mike is transferred to other elements in the scenario. This phenomenon will be of particular interest in a probabilistic setting, since it is a clear effect of coherence that can be captured probabilistically.

Evidential gaps and scenario consequences Finally, the global perspective of a scenario can give us an idea of elements that are clearly missing from the evidence or from the scenario. For example, the scenario about John stabbing Mike includes an element that John grabbed a knife. This element requires evidence to support that a knife was available for John to grab. A scenario-based approach might uncover that such evidence is lacking (or similar evidence stating that John had a knife with him). When a crucial element of a scenario remains unsupported, this is called an evidential gap (Bex, 2011).

Using the coherence of a scenario, one might also 'predict' additional states or events that are to be expected given the scenario. For example, when considering a scenario about John stabbing Mike, the scenario scheme shows that the event 'John ran off' is to be expected in combination with the stabbing. When a scenario 'predicts' that an additional event might have taken place, this is called a scenario consequence (Bex, 2011).

2.4 Summary

In this chapter, we discussed background information on probabilities, Bayesian networks and the scenario-based approach to reasoning with legal evidence. In this section, a brief summary of some of the most relevant notions is provided.

In a probabilistic setting, one is often interested in the posterior probability $\Pr(h|e)$ of a hypothesis h (for example, the suspect committed the crime) given evidence e (the collection of evidence that is available). A Bayesian network can be used to compute such a probability (and many others). A Bayesian network is a compact representation of a joint probability distribution. Nodes in the network represent variables in the domain and the graphical structure of a Bayesian network contains information about (in)dependencies between these variables. Each node has a conditional probability table, which specifies the conditional probability distributions for that node given any configuration of its parents. Constructing a Bayesian network graph and eliciting the required probabilities can be a complex task. Methods exist to assist a modeller in this task.

When using scenarios in reasoning with legal evidence, several alternative scenarios are formed and compared to find the best scenario. This is a form of inference to the best explanation (IBE), in which the best of several alternative explanations is found. A scenario is a coherent collection of states, events and subscenarios. Although there is no clear consensus on what makes a scenario coherent, the view we adopt is to divide coherence into three factors: completeness, consistency and plausibility. When comparing scenarios, one must take into account the quality (in the sense of coherence), as well as the evidential support (or evidential coverage) of a scenario. The latter is crucial to avoid the common pitfall of choosing a good scenario over the true scenario. Finally, deciding between scenarios will always be a trade-off, particularly when deciding between a good scenario with little evidential support and a scenario of lower quality with better evidential support.

The scenario-based approach has several characteristics, most notably that it provides a global perspective on a case. Related to this are the notions of evidential gaps and scenario consequences: since a scenario provides a global view on the case as a whole, it can also point to key pieces of evidence that are missing (evidential gaps) or elements of a scenario that are to be expected (scenario consequences). The scenario-based approach can also be useful to prevent tunnel vision, since it always requires multiple alternative scenarios to be taken into account. Another characteristic of scenarios is that they can be as detailed as desired, making it possible to unfold some event to a high level of detail, while other events can be less detailed. This way, a scenario can be tailored to the specific case and the details required in this case. Finally, a clear effect of the coherence of a scenario is that evidence for some element of the scenario influences our belief in the other elements of the scenario. This phenomenon is called transfer of evidential support.

Part II: Constructing Bayesian networks with scenarios

3. A method for constructing Bayesian networks with scenarios

In this chapter, a method is introduced for constructing a Bayesian network which represents several alternative scenarios, their quality and the evidence related to these scenarios. This method incorporates a combination of a probabilistic approach and a scenario-based approach. By combining the two, we hope to take advantage of the benefits of each separate approach. In particular, the scenario-based approach provides a global perspective on a case whereas a probabilistic approach provides a formal framework for reasoning with uncertainty. While the use of scenarios can help to prevent tunnel vision, the probabilistic approach can help to alleviate the common pitfall that a good scenario is chosen over the true scenario.

As a combination of scenarios and probabilities we propose a Bayesian network which represents the alternative scenarios in a case. This network can then be used to calculate the probability of each scenario given the evidence. To represent scenarios in a Bayesian network, we provide an interpretation of the global coherence of a scenario and its resulting transfer of evidential support in terms of probabilities, as well as an interpretation of scenario quality in terms of completeness, consistency and plausibility in the context of a Bayesian network.

Since the construction of a Bayesian network is not a straightforward task, we furthermore propose a construction method that can be used to gradually construct a Bayesian network representing alternative scenarios, scenario quality and the evidence in a case. This construction method mainly focusses on constructing the graph of a Bayesian network, but it also puts some constraints on the probabilities in the network, needed to represent coherence and plausibility.

In Section 3.3, this construction method is presented, building on the previous sections in which the necessary tools are developed. First, in Section 3.1, four narrative idioms are proposed that can be used to represent scenarios in a Bayesian network. These idioms serve as building blocks for the construction of a Bayesian network graph. Next, in Section 3.2, an interpretation is provided of scenario quality in a Bayesian network context. This section uses the terminology proposed by Pennington and Hastie (1993) and later formalised by Bex (2011), who defined scenario quality in terms of completeness, consistency and plausibility. Together, the idioms and the interpretation of scenario quality serve as a basis for the construction method proposed in Section 3.3. This construction method uses the idea that a scenario can be gradually unfolded to the desired level of detail. In combination with the narrative idioms and the interpretation of scenario quality, this leads to a five-step procedure for constructing a Bayesian network, presented in Section 3.3.

3.1 Narrative idioms

In this section, we propose four narrative idioms, specifically tailored to representing alternative crime scenarios in a Bayesian network. These idioms serve as basic building blocks for the graph of a Bayesian network, inspired by the work of Hepler et al. (2004), who proposed the use of such 'off the shelf', ready-made structures for constructing Bayesian networks for legal evidence, and Fenton et al. (2013) who developed a list of legal idioms. The idioms proposed by Fenton et al. (2013) are very well suited for representing typical situations when working with legal evidence, such as evidence supporting a hypothesis or an alibi. However, the representation of scenarios in a Bayesian network requires a different kind of idiom. Two reasons for this are that firstly, a scenario is concerned with the entire case on a global level, rather than with the details of a case on a local level, and secondly, to properly capture the characteristic properties of a scenario, it is crucial that the coherence of a scenario is taken into account and represented in a Bayesian network.

To cater to these needs of a global perspective and the representation of coherence, the scenario scheme idiom was developed (Section 3.1.1). The scenario scheme idiom is specifically intended to represent a scenario in a Bayesian network graph and thanks to its graphical structure and some constraints on the underlying probabilities, a scenario's coherence is captured. In a similar spirit, the subscenario idiom (Section 3.1.2) can be used to represent coherent substructures, called subscenarios, within a scenario. These substructures can be used to unfold parts of the scenario to add more details, as will be discussed in Section 3.3.2. The variation idiom (Section 3.1.3) is specifically made for representing a scenario in which a small variation of an event also needs to be taken into account, as this situation may be encountered when unfolding a scenario into more details. Finally, a key concept in the scenario-based approach is the comparison of several alternative scenarios. To be able to compare multiple scenarios in one Bayesian network, the merged scenarios idiom (Section 3.1.4) was developed. This idiom can be used once several scenarios have been represented in separate structures, combining these to obtain one Bayesian network for a case.

3.1.1 The scenario scheme idiom

The purpose of the scenario scheme idiom is to enable the representation of a scenario in a Bayesian network as a coherent cluster of states and events. In Section 2.3.1, the coherence of a scenario was ascribed to an underlying pattern or scenario scheme, mapping out the basic structure to which a coherent scenario should adhere. The scenario scheme idiom uses the concept of a scenario scheme to represent a scenario in a Bayesian network. Moreover, as we will show below, the structure of the resulting network is such that it captures the phenomenon of transfer of evidential support, which is an effect of coherence (see Section 2.3).

From scenario schemes to idioms As an example of a scenario, consider a scenario that follows a typical burglary scheme:

Example (A scenario scheme for a burglary). X had a motive to steal. X forced entry into a house. X stole items.

All of the elements in the burglary scheme are propositions that serve as general templates for the actual propositions in a scenario. With a scenario scheme that lays out the structure, a scenario can be formed by filling in the specific propositions of the scenario. With the example burglary scheme, a scenario can be formed by specifying exactly what the motive is (for example, the suspect needed money), how the suspect entered the building (for example, the suspect broke a window) and what items were stolen (for example, the suspect stole a laptop). In what follows, we assume that such scenario schemes exist for various crime scenarios. A collection of scenario schemes can be converted to a collection of idioms, such that a modeller has a collection of scenario scheme idioms to choose from when constructing a network.

To use a scenario scheme as an idiom, it needs to be cast into a Bayesian network structure. Afterwards, the resulting idiom can be used to fill in the elements of a scenario as described above, thereby modelling a scenario in a network structure. However, the elements of a scenario scheme are propositions, while the elements of a Bayesian network are variables. To represent a proposition as a node in a Bayesian network, we use the proposition as the name of the variable, while the possible values of this node are true (`T`) and false (`F`). As a result, the value-assignment `Variable = T` corresponds to the proposition.

Although a scenario scheme is meant to specify the structure of a scenario, the connections between elements of a scenario scheme may remain implicit. In our example of a burglary scheme, no connections were included explicitly, although clearly there are some connections between motive, forced entry and removing items. To translate a scenario scheme to a scenario scheme idiom, these connections need to be made explicit and in particular the (in)dependencies need to be modelled correctly in a Bayesian network. A good heuristic to find the relations between variables is to draw an arrow from one variable to another whenever the first variable can be thought of as a cause of the second variable (Jensen and Nielsen, 2007) (see also Section 2.2). For the purpose of explaining the network at a later stage (see Part III), we furthermore propose to label arrows within the scenario, leading to an annotated Bayesian network structure. Although many different labels could be used, for our current purposes we distinguish between causal (labelled ‘c’) and temporal (labelled ‘t’) connections.

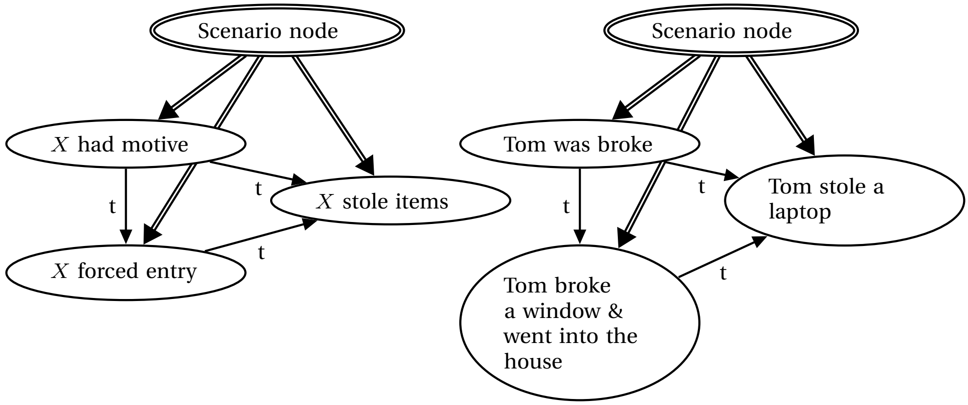
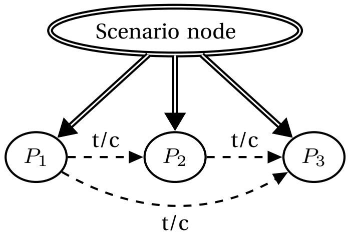


Figure 3.1: The burglary scheme idiom (left) and an instantiation of it (right). Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario.

Capturing coherence When a collection of states and events adheres to a scenario scheme, it forms a coherent scenario. To capture this coherence in a Bayesian network graph, a scenario scheme idiom is structured such that it forms a cluster of nodes. Probabilities are specified such that the elements of a scenario behave as a cluster (more on this will follow below) and visually the graph clearly shows which nodes form a scenario together.

The cluster-like structure of the scenario scheme idiom is due to an additional node called the *scenario node*. This is a boolean node with values true (`T`) and false (`F`) such that `Scenario node = T` represents the situation in which the scenario as a whole is true. Arrows are drawn from the scenario node to each element of the scenario; an example of the typical burglary scenario scheme as a scenario scheme idiom is shown in Figure 3.1 on the left, with an instantiation of this idiom on the right. Double arrows are used to connect the scenario node to each element node, as a visual indication of the underlying constraints on the probabilities representing the coherence of a scenario. Technically these arrows do not differ from other Bayesian network arrows. The underlying constraints are based on the following two principles:

1. whenever the scenario is true, all of its elements must be true; and
2. when the scenario is not true, elements of it might still be true.



Scenario node	T		F	
P_1	T	F	T	F
$P_2 =$ T	1	1	...	...
$P_2 =$ F	0	0	...	...

Figure 3.2: The scenario scheme idiom (left) and an example of a probability table (right) for an element of the scenario. Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario, dashed arrows signify possible connections, depending on the scenario scheme.

To capture this, the probability table of an element node in a scenario is constrained such that any probability conditioned on `Scenario node = T` is equal to 1. In other words, for an element node P_i with parents in the scenario $\text{pa}_S(P_j)$ (that is, parents among other element nodes of the same scenario), we fix the following probabilities for any assignment to $\text{pa}_S(P_i)$:

- $P(P_i \texttt{ = T}|\texttt{Scenario node = T}, \text{pa}_S(P_i)) = 1$
- $P(P_i \texttt{ = F}|\texttt{Scenario node = T}, \text{pa}_S(P_i)) = 0$

The probabilities conditioned on `Scenario node = F` are not constrained. The general structure of a scenario scheme idiom and the constraints on the probability table are shown in Figure 3.2.

With the scenario scheme idiom structured as shown in Figure 3.2, with a scenario node connected to each element node, the phenomenon of transfer of evidential support is captured. Consider the following example:

Example (Transfer of evidential support). Suppose we come home to find a broken window. With this evidence, it not only becomes more probable that someone forced their way into the house, it also becomes more probable that items were stolen.

In the scenario scheme idiom in Figure 3.1, the events 'X forced entry' and 'X stole items' are each connected to the scenario node. The two nodes are connected by a diverging connection via the scenario node and since the scenario node itself is never instantiated, the two elements in the scenario are d-connected and there can be an influence between them. Moreover, due to the constraint on the probabilities, in the absence of other influences, the probability Pr(`Someone took items = T`) increases as the probability Pr(`Someone broke the window = T`) increases as a result of the evidence of a broken window.

The structure of the scenario scheme idiom Summarising the ideas discussed above, we can now define a scenario scheme idiom as follows:

Definition 3.1.1 (Scenario scheme idiom)**.** A *scenario scheme idiom* is a Bayesian network fragment of which the graph $(\mathcal{V}, \mathcal{E})$ and the probabilities are constrained as follows:

- $\mathcal{V}$ consists of a boolean scenario node $\texttt{ScN}$ which represents the scenario as a whole and boolean nodes P which each represent an element of the scenario scheme; and
- $\mathcal{E}$ consists of unlabelled connections $(\texttt{ScN}, P)$ from $\texttt{ScN}$ to each element node P (drawn as double arrows); and
- $\mathcal{E}$ possibly contains labelled connections (P, Q, x) between element nodes P and Q with label x; and
- The probability table for each element node P is constrained such that for any assignment to parent nodes within the same scenario,

$$\Pr(\texttt{P = T}|\texttt{ScN = T}, \mathbf{pa}_S(P_i)) = 1.$$

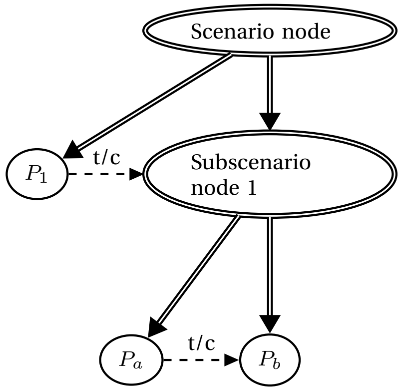
3.1.2 The subscenario idiom

The elements of a scenario may be organized in subscenarios: coherent substructures within the scenario. Consider the following example about a burglary:

Example (A scenario with a subscenario). Tom was broke. A loose brick was near a house. Tom threw the brick at the window. The window broke. Tom went into the house. Tom stole a laptop.

Within this scenario, the following elements together form a subscenario about how Tom broke a window and went into the house: 'A loose brick was near the house. Tom threw the brick at the window. The window broke. Tom went into the house'.

The structure of the subscenario idiom A subscenario is a coherent substructure within the scenario. To model an instance of a subscenario, a scenario scheme (about someone breaking a window, for instance) can be used as a template for the coherent structure of that subscenario. To capture the coherence of a subscenario within a scenario in a Bayesian network, we apply a similar structure as with the scenario scheme idiom. Again, the elements that are jointly coherent are clustered using an additional node representing the subscenario as a whole: the subscenario node. Arrows point from the subscenario node to each element of the subscenario. And since the subscenario is itself an element of the scenario, the subscenario node



Subscenario node	T		F	
P_a	T	F	T	F
$P_b = \mathtt{T}$	1	1	...	...
$P_b = \mathtt{F}$	0	0	...	...

Figure 3.3: The subscenario idiom as part of a scenario scheme idiom and an example of a probability table for an element of the subscenario. Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario, dashed arrows signify possible connections.

is an element node in the scenario. The structure of the subscenario idiom is shown in Figure 3.3. Again, the probability tables of the elements of the subscenario are constrained, as shown in the table in Figure 3.3, based on the same principles as used in the scenario scheme idiom. Double arrows signify the special connections between the subscenario node and the elements of the subscenario as a result of these constrained probabilities.

This leads to the following definition of the subscenario idiom:

Definition 3.1.2 (Subscenario idiom)**.** A *subscenario idiom* is a scenario scheme idiom for which

- the scenario node is now called a subscenario node; and
- the subscenario node is an element in some scenario (a child of some scenario node).

In Figure 3.4, the example burglary scenario is represented with the subscenario idiom. The node ‘Tom broke a window & went into the house’ is the subscenario node that represents the subscenario as a whole. It also serves as an element node in the global scenario.

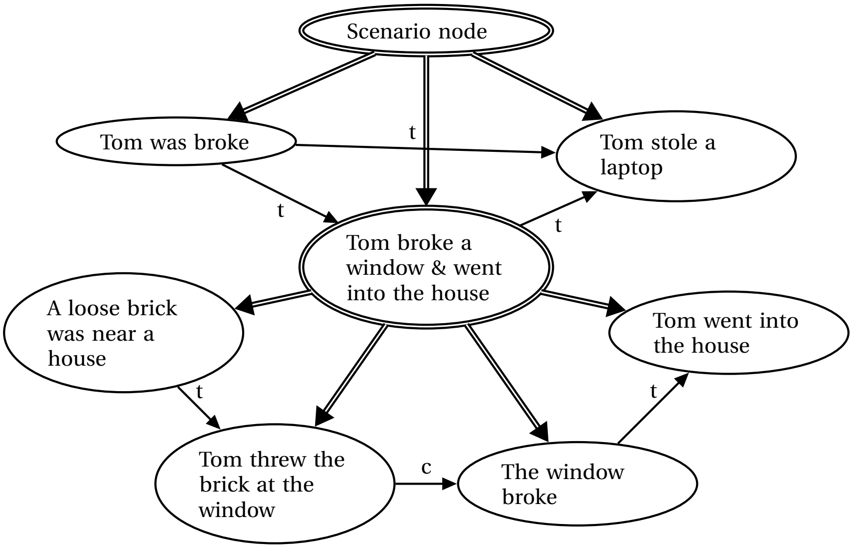


Figure 3.4: A burglary scenario with subscenario. Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario.

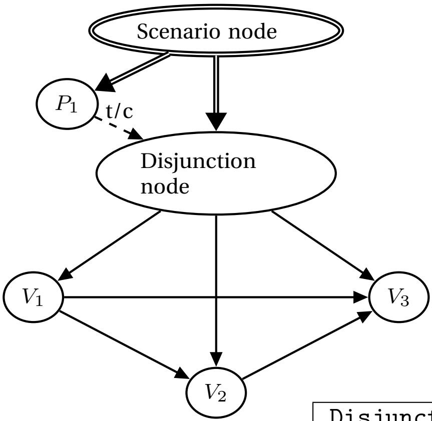
3.1.3 The variation idiom

Suppose the example scenario about a burglary needs to be modelled, but it is unclear whether Tom entered the house by breaking a window, forcing a door or picking a lock. The scenario that needs to be modelled is thus as follows:

Example (A scenario with variations). Tom was broke. Tom broke a window/forced a door/picked a lock and went into the house. Tom stole a laptop.

When a scenario includes several mutually exclusive alternatives for an element (or subscenario) in that scenario, these are said to be variations within the scenario. To model the above example, one option is to model three scenarios, one in which Tom broke a window, one in which he forced a door and one in which he picked a lock. However, this leads to a complex graphical structure since much of the structure will have to be unnecessarily duplicated. Instead, we propose the variation idiom, a structure that can be used to represent such variations within a scenario. The idiom is particularly intended for mutually exclusive variations that do not influence the overall conclusion of the scenario (namely that it was Tom who broke in).

Why a variation idiom is needed Although in a typical Bayesian network modelling approach one might model a variation as alternative values of a single node (a



Disjunction node	T	F
V_1 = T	0.33	0
V_1 = F	0.67	1

Disjunction node	T		F	
V_1	T	F	T	F
V_2 = T	0	0.5	0	0
V_2 = F	1	0.5	1	1

Disjunction node	T				F			
V_1	T		F		T		F	
V_2	T	F	T	F	T	F	T	F
V_3 = T	0	0	0	1	0	0	0	0
V_3 = F	1	1	1	0	1	1	1	1

Figure 3.5: The variation idiom (left) and an example of the probability tables (right) for the case with three variations. Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario. Connections between variations are needed to make sure that exactly one of the variations holds if the disjunction node is true.

node 'forced entry' with values 'smashed window', 'forced door' and 'picked lock'), in our scenario context a different approach is called for. While we do want to specify entirely new scenarios for each variation as described above, it is desirable to represent each variation as a separate node, making it possible to have subscenarios as variations. This is because a separate node for each variation can now serve as a subscenario node with elements of the subscenario as children. This is impossible when the variations are alternative values of a single node. The variation idiom with three variations is shown in Figure 3.5 (left), including an example of the probability tables for a case with three variations (right). These probability tables are partially constrained, as will be further explained below. An example of the burglary scenario with a subscenario among its variations is shown in Figure 3.6.

The structure of the variation idiom The variation idiom (see Figure 3.5 for the case with three variations) models each variation as a node V_i, assuming that the variations are ordered as an enumerated list. Furthermore, an additional node called the disjunction node represents the entire disjunction, namely that exactly

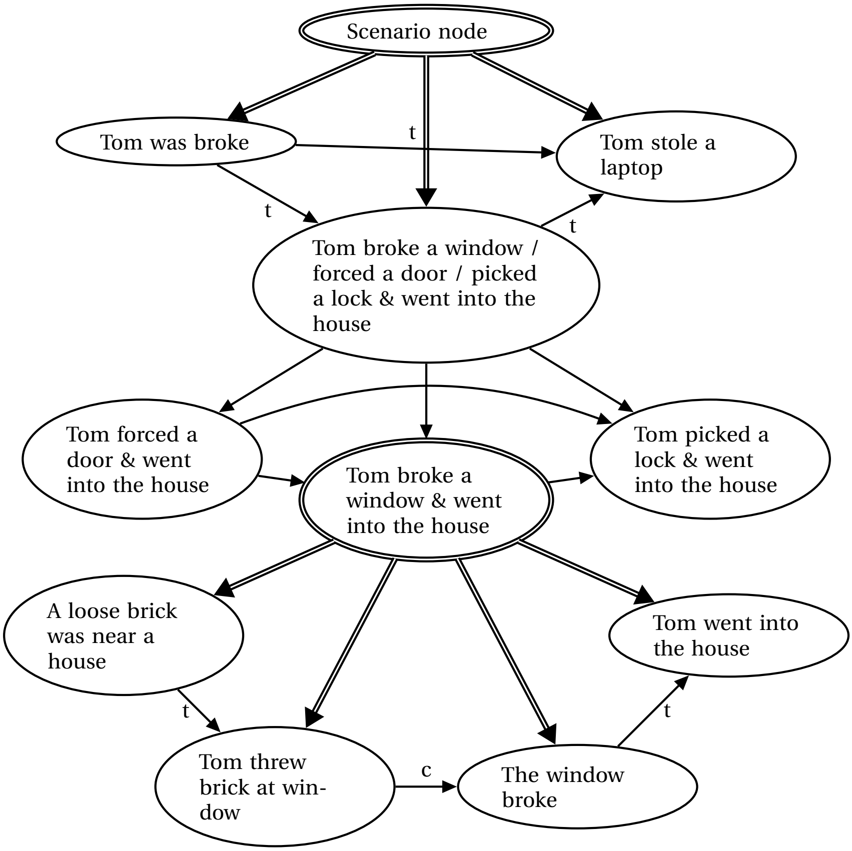


Figure 3.6: A burglary scenario with variations and a subscenario. Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario. The node 'Tom broke a window & went into the house' is a variation and a subscenario node.

one of the variations must hold. This disjunction node also serves as an element node in the scenario. Arrows point from the disjunction node to each variation V_i and for any pair of variations V_i and V_j there is an arrow $V_i \to V_j$ if and only if $i < j$. Since arrows point only from the disjunction node to a variation and from a variation with a lower index to a variation with a higher index, the result will be acyclic. Connecting arrows between any pair of variations are needed to be able to express that if the disjunction node is true, exactly one variation V_i must be true.

We thus have the following definition of the variation idiom:

Definition 3.1.3 (Variation idiom)**.** A *variation idiom* is a Bayesian network fragment with a graph $(\mathcal{V}, \mathcal{E})$ and probabilities for which

- $\mathcal{V}$ consists of a boolean disjunction node `disj` which represents the collection of variations and is an element of some scenario (a child of some scenario node); and
- $\mathcal{V}$ consists of boolean variation nodes V_i for each variation; and
- $\mathcal{E}$ consists of unlabelled connections $(\texttt{disj}, V_i)$ from `disj` to each V_i. For each pair of variations V_i and V_j there is a connection (V_i, V_j) if and only if $i < j$; and
- The probability table for each variation node V_i contains probabilities conditioned on `disj` and on all $V_1, ...V_{i-1}$. It is constrained such that

$$\Pr(V_i \texttt{ = T}|\texttt{disj = T}, V_1\texttt{= F}, \ldots, V_j\texttt{= T}, \ldots, V_{i-1}\texttt{= F}) = 0$$

$$\Pr(V_i \texttt{ = T}|\texttt{disj = F}, V_1, \ldots, V_{i-1}) = 0.$$

- In addition, the probability table for the last variation node V_n is constrained such that:

$$\Pr(V_n \texttt{ = T}|\texttt{disj = T}, V_1\texttt{= F}, \ldots, V_{n-1}\texttt{= F}) = 1.$$

In this idiom, the probability tables for variation nodes are constrained to represent that exactly one variation must be true if the disjunction node is true. The ideas behind these probabilities are explained below.

The probabilities The probability tables of the variation nodes are specified as follows to represent that exactly one variation holds if the disjunction node holds:

- If the disjunction node is false, the disjunction as a whole is false so none of the variations can be true. This leads to the following probability for any V_i and any assignment of $V_1, \ldots, V_{i-1}$: $\Pr(V_i \texttt{ = T}|\texttt{disjunction node = F}, V_1, \ldots, V_{i-1}) = 0$;
- If the disjunction node is true, this means that exactly one of the variations must hold.
  - To make sure that at least one variation holds, the last variation V_n (out of n variations) will be true with probability 1 if all other variations are false. Note that for V_n, all other variations V_i with $i < n$ are parents. Therefore, in the probability table it can be specified that $\Pr(V_n \texttt{ = T}|\texttt{disj = T}, V_1\texttt{= F}, \ldots, V_{n-1}\texttt{ = F}) = 1$.

– To make sure that at most one variation holds, for a variation V_i, if there is some V_j= T with $j < i$, then V_i must be false. Therefore, $\Pr(V_i \texttt{ = T}|\texttt{disj = T}, V_1\texttt{= F}, \ldots, V_j\texttt{= T}, \ldots, V_{i-1}\texttt{= F}) = 0$.
– Other probabilities need to be specified by the modeller. In the absence of further information, the variations can be modelled as equally likely.[1] This can be done by setting the probability of a variation V_i to be true, given that all variations with a lower index V_j with $j < i$ are false, to $1/(n-i+1)$. So $\Pr(V_i \texttt{ = T}|\texttt{disj = T}, V_1\texttt{= F}, \ldots, V_{i-1}\texttt{= F}) = \frac{1}{n-i+1}$.

An example of these probabilities for the situation with three variations is shown in the tables in Figure 3.5. Note that the probability table also contains numbers for $P(V_2 = \texttt{T}|\texttt{disj = F}, V_1 = \texttt{T})$, expressing the probability that variation 2 occurs given that the disjunction is false but variation 1 is true, which should be impossible since the disjunction being false would yield variation 1 to be false as well. This situation will indeed never occur, since the probability table for variation 1 guarantees that variation 1 is false whenever the disjunction node does not hold. Hence, these numbers in the probability table for variation 2 are really undefined, but need to be given some value to enable calculations in the Bayesian network. These values can be arbitrary since they will be summed out during the calculations.

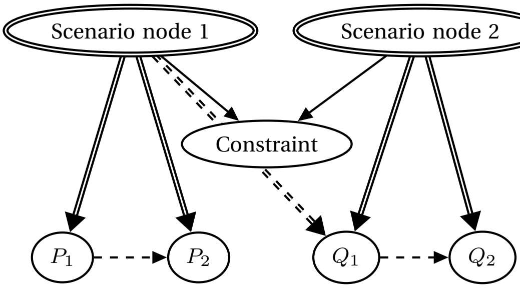
Finally, the probability table for the disjunction node is constrained as any other element in the scenario. The unconstrained numbers should specify the probability of the disjunction (for example, Tom broke a window/forced a door/picked a lock & went into the house) conditioned on the scenario node being false and possibly other elements in the scenario.

Why a different structure is not suitable Note that in the structure of the variation idiom, the arrows from the disjunction node to each variation node V_i are directed from disjunction to variation. A different structure, with arrows from variation to disjunction, may seem more intuitive, but would lead to a conflict with the scenario scheme idiom of which the disjunction node is a part. In particular, the disjunction node would then have incoming arrows from the scenario node and all of the variations, resulting in the following probability (amongst others) that needs to be specified in the conditional probability table of the disjunction node:

$$P(\texttt{disj = T}|\texttt{Scenario Node = true}, V_1 \texttt{ = F}, ..., V_n \texttt{ = F}).$$

By definition of the scenario scheme idiom, this probability needs to be 1, while for the variation idiom to work properly it would need to be 0. To see the latter, recall that when all variations are false, the disjunction node should also be false. Therefore, a structure with arrows from variations to the disjunction node cannot capture the variations as desired while leaving the scenario idiom intact.

[1] Note that specifying numbers in the absence of information runs the risk of falsely suggesting that these exact numbers are known.



Scenario node 1	T		F	
Scenario node 2	T	F	T	F
Constraint = s1	0	1	0	0.5
Constraint = s2	0	0	1	0.5
Constraint = NA	1	0	0	0

Figure 3.7: The merged scenarios idiom (left) and the probability table for the constraint node in the situation with two scenarios (right). Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario. The dashed double arrow from Scenario node 1 to element P_b of scenario 2 suggests that an element can be part of multiple scenarios.

3.1.4 Merged scenarios idiom

By representing multiple scenarios in one Bayesian network, various scenarios can be compared. In order to model all scenarios in one network, the merged scenarios idiom can be used. This idiom puts a constraint on the scenario nodes of separate scenarios, making sure that at most one scenario can be true.

Mutually exclusive scenarios The merged scenarios idiom assumes that all scenarios in it are mutually exclusive. This means that the merged scenarios idiom should only be applied to scenarios for which this is the case. Consider the following example:

Example (Alternative scenarios). In a burglary case, the police finds fingerprints of Sylvia and footsteps of Tom. Three alternative scenarios could be the following:

1. Sylvia and Tom committed the burglary
2. Sylvia was at the crime scene on an unrelated occasion
3. Tom committed the burglary alone.

Note that scenarios (2) and (3) can be the case simultaneously. In fact, together they account for the evidence that was found. In this situation, the merged scenarios idiom should only be applied to scenarios (1) and (3), since they are alternative scenarios about someone committing a crime. Scenario (2) can still be modelled in the network as an alternative explanation of the evidence, but will not be connected to the constraint node.

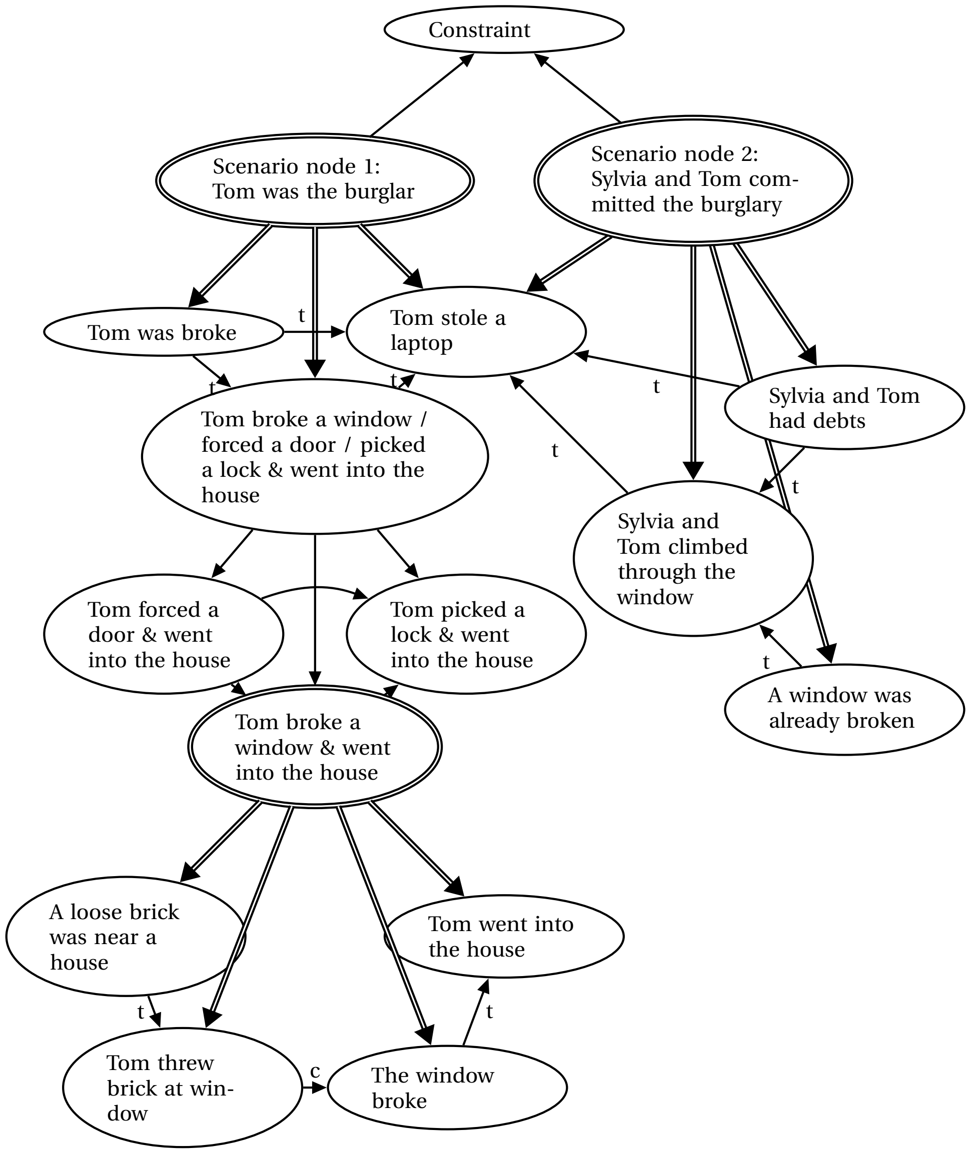


Figure 3.8: An example of two scenarios merged in one network. Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario. The event 'Tom stole a laptop' is an element of both scenarios.

The structure of the merged scenarios idiom With the merged scenarios idiom, two or more representations of scenarios are combined to form one network. A constraint node is connected to all mutually exclusive scenario nodes as shown in Figure 3.7 (left): for any scenario node $\texttt{ScN}_i$ in the collection of n scenarios $\{\texttt{ScN}_1, \ldots, \texttt{ScN}_n\}$ there is an arrow from $\texttt{ScN}_i$ to the constraint node C. An example of this idiom applied to two burglary scenarios is shown in Figure 3.8.

The constraint node has values s_i for each scenario $\texttt{ScN}_i$ and one value NA (for not applicable) to denote that an illegal combination of nodes is considered (see Fenton et al. (2011) for this constraint node construction). The probability table of the constraint node is shown in Figure 3.7 and it is such that when multiple scenarios are simultaneously true, the constraint node will have value NA. Evidence for the constraint node is now entered such that the value NA has a probability of 0 to occur. This is so-called 'soft evidence' (or virtual evidence), which is not an observation of the value of the node, but rather a 'partial' observation about which value the node cannot take. Other values now have a summed probability of 1. Setting such soft evidence ensures that multiple scenarios cannot be true simultaneously.

The definition of the merged scenarios idiom is as follows:

Definition 3.1.4 (Merged scenarios idiom)**.** A *merged scenarios idiom* is a Bayesian network fragment such that the graph $(\mathcal{V}, \mathcal{E})$ and probabilities are as follows:

- $\mathcal{V}$ consists of boolean scenario nodes $\texttt{ScN}_i$ with $1 \leq i \leq n$ which each represents a scenario and a constraint node `Constraint` with values s_i for each scenario node and a value `NA`; and

- $\mathcal{E}$ consists of unlabelled connections $(\texttt{ScN}_i, \texttt{Constraint})$ from each scenario node $\texttt{ScN}_i$ to the constraint node; and

- The probabilities are constrained such that

$$\Pr(\texttt{Constraint = } s_i | \texttt{ScN}_1 \texttt{ = F}, \ldots, \texttt{ScN}_i \texttt{ = T}, \ldots, \texttt{ScN}_n \texttt{ = F}) = 1$$

$$\Pr(\texttt{Constraint = NA} | \texttt{ScN}_1 \texttt{ = F}, \ldots, \texttt{ScN}_i \texttt{ = T}, \texttt{ScN}_j \texttt{ = T}, \ldots, \texttt{ScN}_n \texttt{ = F}) = 1.$$

- Soft evidence is entered in the network such that $\Pr(\texttt{Constraint = NA}) = 0$.

The probabilities of the constraint node are thus constrained such that if $\texttt{ScN}_i$ has value `T`, then the constraint node has value $\texttt{s}_i$ with probability 1, unless other scenario nodes also have value true, in which case the constraint node has value `NA` with probability 1.

After merging When merging two scenario structures in the merged scenarios idiom, an element of one scenario may also be an element of another scenario. For example, the scenarios about Tom and Sylvia both contain the element `Tom stole a laptop`. When this happens, the separate nodes that occurred in different scenarios should be replaced by a single node with all arrows pointing to and from the original nodes now pointing to and from the single new node. As can be seen in Figure 3.8, this results in a structure in which one node is a child node to two different scenario nodes.

3.2 Representing scenario quality

In this section we investigate how the quality of a scenario can be captured in a Bayesian network that is built with the idioms from Section 3.1. Representing the quality of a scenario in a Bayesian network is a crucial step in embedding the scenario-based approach in our method, since the scenario-based approach relies on comparing alternative scenarios. When comparing scenarios, the evidence supporting each scenario needs be taken into account, but also the quality of each scenario. Together, these two factors determine which scenario is the most acceptable. A Bayesian network is of itself already suitable for drawing conclusions about the evidence supporting each scenario, since the effect of instantiating evidence on the probability of each scenario node can be read from the network directly. However, thinking about the quality of a scenario in a Bayesian network context is not straightforward.

In the scenario-based approach, the quality of a scenario can be expressed in terms of the following three factors: completeness, consistency and plausibility (see Section 2.3 for background on these concepts). In the sections below, each of these factors is interpreted in a Bayesian network context. The purpose of providing a Bayesian network interpretation of these terms is twofold: firstly, it makes it possible to represent the quality of a scenario in a Bayesian network and secondly, the same terminology makes it possible to explain the results of a Bayesian network in these notions from the scenario-based approach. The latter will be further explored in Part III of this thesis.

3.2.1 Completeness

According to Pennington and Hastie (1993), a scenario is complete when it has all its parts. This notion was formalized by Bex (2011) in the context of his Hybrid Theory, using scenario schemes: a scenario is complete when it fits and completes a scenario scheme (see also Section 2.3). Such scenario schemes were translated to scenario scheme idioms in Section 3.1, so inspired by Bex's definitions, we now define the concept of completeness in the context of Bayesian networks, using scenario scheme idioms as follows:

Definition 3.2.1 (Fitting a scenario scheme idiom)**.** A scenario *fits* with a scenario scheme idiom when for every proposition p in the scenario, there is a corresponding node P in the scenario scheme idiom.

Definition 3.2.2 (Completing a scenario scheme idiom)**.** A scenario *completes* a scenario scheme idiom when for every node P in the scenario scheme idiom there is some corresponding proposition p in the scenario.

This results in the following definition of completeness, relative to scenario scheme idioms:

Definition 3.2.3 (Completeness)**.** A scenario is *complete* when it fits and completes a scenario scheme idiom.

Completeness in the modelling process In practice, a finished Bayesian network should only model complete scenarios. To this end, when an incomplete scenario is encountered, the modeller might 'fill in the gaps' of a scenario by adding events such that a scenario completes an appropriate scenario scheme idiom.

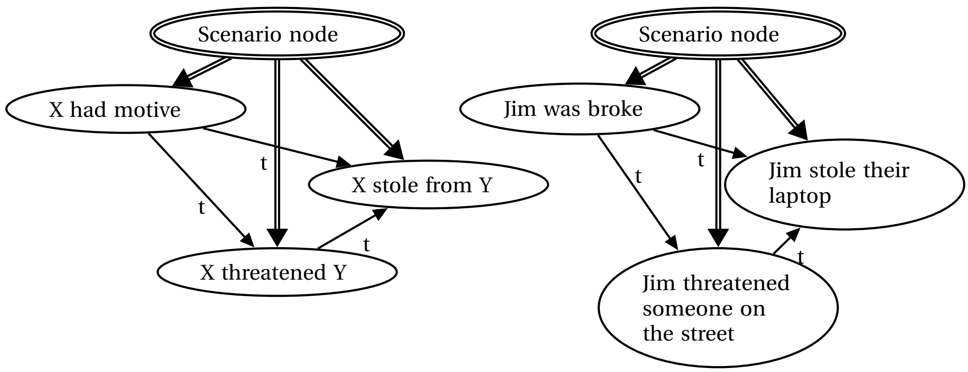


Figure 3.9: A scenario scheme idiom for a robbery (left) and a scenario about Jim robbing someone (right).

As an example of an incomplete scenario that might be encountered during the modelling process, consider the following:

Example (An incomplete scenario). Jim was broke. Jim stole a laptop.

This incomplete scenario fits the scenario scheme idiom for a robbery (see Figure 3.9 (left)), but it lacks an element about how Jim threatened a person before stealing their laptop. Completing the scenario with respect to this scheme results in a graph as shown in Figure 3.9 (right).

Note that the same incomplete scenario could be completed differently when a different scenario scheme idiom is used. For example, the same incomplete scenario also fits the scenario scheme from Figure 3.1 about a burglary, but in that case it lacks an element about how Jim entered the house. By adding the elements to complete the scenario relative to this scenario scheme idiom, the modeller can model this as a second scenario. Note that the modeller can model as many completed scenario as are deemed relevant for the case; they are simply modelled as alternative possibilities and the model will be used to then find which scenario is most probable.

3.2.2 Consistency

When a scenario is inconsistent, this is because two or more elements of the scenario are jointly inconsistent (Wagenaar et al., 1993; Pennington and Hastie, 1993; Bex, 2011). This can be represented in the network by adding a constraint node construction similar to the the constraint in the merged scenarios idiom from Section 3.1. A constraint node is added and connected to the nodes that are jointly inconsistent. The probabilities are set such that the elements that are jointly inconsistent ($P_1, P_2...$) have a probability of 0 to be true at the same time ($\Pr(P_1 = \mathtt{T}, P_2 = \mathtt{T}, ...) = 0$). Consistency and inconsistency can thus be defined as follows:

Definition 3.2.4 (Inconsistency)**.** A scenario is *inconsistent* when there is a set of nodes $\{P_1, ...P_n\}$ with $n \geq 2$ corresponding to propositions in the scenario for which $\Pr(P_1 = \mathtt{T}, ..., P_n = \mathtt{T}) = 0$.

Definition 3.2.5 (Consistency)**.** A scenario is *consistent* when it is not inconsistent.

As an example of an inconsistency in a scenario, consider the following:

Example (An inconsistent scenario). Suppose a scenario contains the elements 'Jim was at the cinema at 8 p.m.' and 'Jim robbed someone at 8 p.m.'. These two elements are jointly inconsistent, which makes the scenario inconsistent.

To represent this inconsistency, a constraint node is added to the network and arrows are drawn from nodes `Jim was at the cinema at 8 p.m.` and `Jim robbed someone at 8 p.m.` to the constraint node. When a scenario's inconsistency is modelled with a constraint, it results in a probability of 0 for the scenario node of that scenario. To see this, consider a scenario with two inconsistent elements P_1 and P_2. If the scenario node had value `true`, these elements P_1 and P_2 would also have value `true` with probability 1. Since the constraint forces that they cannot both be true, it follows that the scenario node must have value `false` with probability 1. A scenario which contains an inconsistency is thus not considered a viable alternative.

3.2.3 Plausibility

The plausibility of a scenario depends on how well the scenario matches our common sense knowledge of the world (Wagenaar et al., 1993; Pennington and Hastie, 1993; Bex, 2011). In our Bayesian network context, the common sense knowledge is arguably in the prior probabilities of the network: if our a priori degree of belief in an element is high, this element is close to our common sense knowledge of the world. When the a priori degree of belief in an element is low, we find it less credible based on our common sense knowledge. This leads to the following definitions of the plausibility of elements:

Definition 3.2.6 (Plausibility of an element)**.** The *plausibility of an element* P in a scenario, is given by $\Pr(P = \texttt{T})$.

The plausibility of an element is thus given by the marginal probability of the corresponding variable, which is the probability of that variable unconditioned on any other variables. Calculating the conditional probabilities required for the probability table when given the marginal probability of a variable is not straightforward; a strategy for representing plausibility correctly in the probability tables of a network will be presented below.

When an element P has one or more parents within the scenario, one can also consider the plausibility of P given parents $Q_1, \dots Q_n$, which is given by $\Pr(P = \texttt{T} | Q_1 = \texttt{T}, \dots, Q_n = \texttt{T})$. And since the scenario node represents the scenario as a whole, we can also define the plausibility of the scenario:

Definition 3.2.7 (Plausibility of a scenario)**.** The *plausibility of a scenario*, which is represented by scenario node `ScN`, is given by $\Pr(\texttt{ScN = T})$.

As a rule of thumb the plausibility of a scenario should always be smaller than the plausibility of its least plausible element. This is needed to be able to correctly model the desired plausibility of elements in a Bayesian network, as will become apparent from the calculations below (in the paragraphs about representing plausibility). From this, it may seem intuitive that the plausibility of a scenario with less plausible elements should be lower than that of a scenario with more plausible elements, but this is not a technical requirement.

An example of a scenario with an implausible element is the following:

Example (A scenario with an implausible element). The second scenario in Figure 3.8 contains the element 'A window was already broken'. This is an implausible element.

The implausibility of this element will be reflected in the network with a low probability $\Pr(\texttt{A window was already broken = T})$.

Representing the plausibility of elements To represent the plausibility of an element in a Bayesian network, the probability table of that element needs to be elicited accordingly. By first establishing how plausible the modeller finds a certain element of a scenario, the required numbers for the probability tables can be calculated. For instance, suppose a modeller wants to model the implausibility of 'A window was already broken' by setting the following low probability: $\Pr(\texttt{A window was already broken = T}) = 0.01$. This probability is not in any of the probability tables, but it will be calculated as follows (with the scenario node abbreviated to ScN and in the absence of other parents):

$$\begin{aligned}
&\Pr(\texttt{A window was already broken = T})\\
&= \Pr(\texttt{A window was already broken = T}|\texttt{ScN = T}) \cdot \Pr(\texttt{ScN = T})\\
&+ \Pr(\texttt{A window was already broken = T}|\texttt{ScN = F}) \cdot \Pr(\texttt{ScN = F}).
\end{aligned}$$

Since $\Pr(\texttt{A window was already broken = T}|\texttt{ScN = T}) = 1$ by definition of the scenario scheme idiom, the probability that should go in the probability table, namely $\Pr(\texttt{A window was already broken = T}|\texttt{ScN = F})$, can be calculated based on $\Pr(\texttt{A window was already broken = T})$ as soon as the probability table of the scenario node has been elicited, thereby specifying the probability $\Pr(\texttt{ScN = T})$. The probability $\Pr(\texttt{A window was already broken = T} \mid \texttt{ScN = F})$ that needs to be elicited for the probability table, to properly model the implausibility, is then found as follows:

$$\begin{aligned}
&\Pr(\texttt{A window was already broken = T}|\texttt{ScN = F})\\
&= \frac{\Pr(\texttt{A window was already broken}) - \Pr(\texttt{ScN = T})}{\Pr(\texttt{ScN = F})}\\
&= \frac{0.01 - \Pr(\texttt{ScN = T})}{\Pr(\texttt{ScN = F})}\\
&= \frac{0.01 - \Pr(\texttt{ScN = T})}{1 - \Pr(\texttt{ScN = T})}.
\end{aligned}$$

These calculations show that the plausibility of a scenario (given by the probability that the scenario node is true) can never be larger than the plausibility of any element in that scenario, since otherwise a negative number would have to be entered in the probability table of that element.

Representing the plausibility of elements with parents When a node has multiple parents, one can represent the plausibility of that element given its parents. Consider, for example, the connection between events 'A bird flew into the window' and 'The window broke'. The element 'The window broke' is can be considered implausible given the element 'A bird flew into the window', since birds do not usually break windows. To model this, an appropriate probability needs to be set for $\Pr(\texttt{The window broke = T}|\texttt{A bird flew into the window = T}, \texttt{Scenario node = F})$ in the probability table to represent this plausibility. Again, this can be calculated from the desired value for $\Pr(\texttt{The window broke = T}|\texttt{A bird flew into the window = T})$, once the probabilities are known for `A bird flew into the window`. The calculation is then done as follows (with `A bird flew into the window` abbreviated to `A bird...`):

$$\Pr(\texttt{The window broke = T}|\texttt{A bird flew into the window = T}, \texttt{ScN = F}) = \frac{\Pr(\texttt{The window broke = T}|\texttt{A bird...= T}) \cdot \Pr(\texttt{A bird...= T}) - \Pr(\texttt{ScN = T})}{\Pr(\texttt{A bird...= T}|\texttt{ScN = F}) \cdot \Pr(\texttt{ScN = F})}.$$

In sum, the modeller can go through the network starting at the root nodes to fill in the probability tables to represent plausibility. For a node P_1 with only the scenario node as a parent, the plausibility $\Pr(P_1 = \texttt{T})$ can be used to calculate $\Pr(P_1 = \texttt{T}|\texttt{ScN = F})$, as soon as the probabilities of `ScN` have been specified. Similarly, after specifying the probability table of P_1, a node P_2 connected to P_1 can be specified using the plausibility of P_2 given P_1. This way, the concept of plausibility is represented in the network.

3.3 A construction method

The construction of a Bayesian network is not a straightforward task. Many variables may be of interest and eliciting the structure of the graph and the underlying probabilities can be daunting. However, the construction of a Bayesian network can benefit from the scenario-based approach since scenarios provide the context needed to determine which variables are relevant for the specific case at hand. This is especially helpful when modelling legal evidence, since legal cases are often very 'open world' problems in which a diversity of variables may turn out to be relevant, including, for example, the suspect buying a movie ticket three weeks before. By forming several alternative scenarios, a selection is made of contextual variables that may be relevant to the case.

Several alternative scenarios can thus be used as a guideline for constructing a Bayesian network. With the materials from the previous sections, these scenarios and their quality can be represented in a network. In this section, we use these ideas to formulate a step-by-step construction method in which the Bayesian network representing several alternative scenarios is constructed incrementally. This method assumes that a number of relevant scenarios are known, such that the problem that is being addressed is how a Bayesian network can be constructed on the basis of these scenarios. To incrementally construct a Bayesian network for a case, we first propose the technique of unfolding a scenario in Section 3.3.1. This technique is then used in Section 3.3.2 to formulate a procedure consisting of five steps, with which a network can be constructed.

3.3.1 Unfolding a scenario

A legal case may require very specific details about one part of the case and much less detail about other parts. In the burglary case the exact events surrounding the break-in itself deserve a lot of attention, including, for example, the exact order in which the burglar entered different rooms in the house. However, the burglar could probably go into a lot of detail about why he decided to break in, which might not be relevant for the case. Legal cases thus require a scenario about the crime that has different levels of detail for the various events. Scenarios have the property that elements can be unfolded to various levels of detail (see Section 2.3). In daily life, we employ this property often; for example, when telling a friend about our restaurant visit last night, we would include much detail about the food and much less about asking for the check, paying and leaving (Schank and Abelson, 1977). However, if we were to report to the police about this restaurant visit because some money was stolen, we would focus on the payment rather than the quality of the food. Wagenaar et al. (1993) argue in their Anchored Narratives Theory that a story about a crime should be made more precise whenever more details are needed.

With the method of unfolding a scenario, a Bayesian network for a case is gradually constructed, employing the narrative property that various parts of a scenario can be elaborated upon at various levels of detail. Starting with an initial scenario about the case, questions are asked to determine step by step whether more detail is needed about elements of this scenario. When more details are required, the element node from the initial scenario is *unfolded* to a subscenario.

Gradually constructing a network When an initial scenario has been modelled, an event in that scenario can be unfolded to a subscenario by transforming the element node to a subscenario node and adding the elements of the subscenario as children. The transformation from an element node to a subscenario node has no technical implications: all arrows that were connected to this element node remain connected to the subscenario node and the conditional probability table for this node remains the same. As an example of unfolding consider the following:

Example (Unfolding an element in a scenario). An initial version of a burglary scenario is shown in Figure 3.10, on the left. Unfolding the element `Tom broke a window & went into the house` leads to a structure as shown on the right. The element node is now replaced by a subscenario node, with elements of that subscenario providing a more detailed account of the original event.

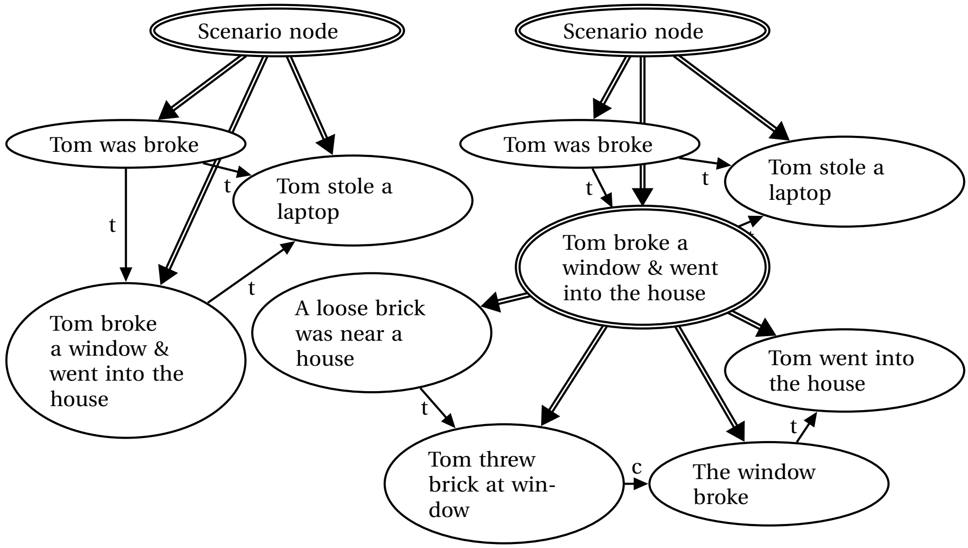


Figure 3.10: Unfolding an initial scenario about a burglary (left) to a scenario with a subscenario (right). Double arrows signify that the underlying probabilities are constrained to represent the coherence of a scenario.

When to unfold As said, the core idea of unfolding scenarios is that some, but not all, elements of a scenario require more detail. Whether an element of a scenario requires unfolding depends on its connections to (possible) evidence and whether adding more detail makes these connections more insightful. For example, the original event `Tom broke a window & went into the house` can be connected to evidence of a witness who saw Tom throwing a brick, but by unfolding the event it can be connected to `Tom threw a brick at window` instead, resulting in a more direct connection. Therefore, when modeling a case, the following three questions can be asked to determine whether an element of a scenario requires unfolding:

1. Is there evidence that can be connected directly to the element node? If so, no unfolding is required.

2. Is there relevant evidence for details of a subscenario for this element? If so, unfolding is required.

3. Would it be possible to find relevant evidence for details of the subscenario for this element? If so, unfolding is required.

By asking the three question above, elements of a scenario are unfolded whenever they lead to relevant evidence being included in the network, or if there is no evidence yet, the possibility of more evidence being found (question 3).

3.3.2 The construction method

Using the narrative idioms, the interpretation of scenario quality in a Bayesian network and the technique of unfolding, a Bayesian network can be constructed with the following procedure (each step will be explained further below):

1. **Represent**: for each scenario, select one or more scenario scheme idioms that the scenario fits. Represent completeness and consistency: if the scenario is incomplete with respect to a scenario scheme idiom, extend the scenario such that it completes the scheme and if the scenario is inconsistent, add a constraint node to model the inconsistency;

2. **Unfold**: for each scenario, unfold to more detailed subscenarios if needed by repeatedly asking the three unfolding questions. Use the subscenario idiom to model the unfolding subscenarios and the variation idiom whenever a variation is encountered. The process of unfolding is finished when the three questions indicate that no more relevant evidence can be added to the structure;

3. **Merge**: use the merged scenarios idiom to merge the scenario structures constructed in the previous step;

4. **Include evidence nodes**: for each piece of evidence that is available, use the idioms from Section 2.2 to represent them as nodes. Additionally, include nodes for evidential data that is to be expected as an effect of elements in the structure.

5. **Specify probability**: specify probabilities as dictated by each idiom and represent plausibility in the probability tables. Elicit other probabilities using elicitation techniques.

This procedure uses the narrative idioms from Section 3.1 to represent each scenario in an initial Bayesian network structure, representing completeness and consistency properties of the scenario (step 1). This initial scenario is then unfolded to the desired level of detail using the unfolding technique from Section 3.3.1, to incrementally construct a network structure (step 2). Several structures representing the alternative scenarios are then merged using the merged scenarios idiom from Section 3.1 (step 3) and evidential nodes are added to the structure using the idioms proposed by Fenton et al. (2013) (step 4). Finally, probabilities are partially specified by the idioms and otherwise used to represent plausibility. Existing elicitation techniques (see Section 2.2) can be used for any remaining probabilities.

3.4 Conclusion

In this chapter, a method was proposed for constructing a Bayesian network. The result is a network which models several alternative scenarios, the quality of these scenarios and the evidence connected to them. This method uses the narrative idioms and the interpretation of scenario quality to construct a Bayesian network for a case.

In Section 3.1, four narrative idioms were proposed for capturing scenarios in a Bayesian network. The scenario scheme idiom was specifically intended for capturing a scenario and its coherence properties in a network structure. The structure of the scenario scheme idiom is such that a scenario is clearly visible in a network as a coherent cluster of events and the probabilities are such that the so-called transfer of evidential support as a consequence of coherence is captured in the network.

The scenario scheme idiom builds on the idea of scenario schemes as underlying structures for a scenario. Each scenario scheme (for example, a typical burglary, or murder, or robbery) is translated to a specific scenario scheme idiom, in such a way that there is a clear correspondence between the resulting structure and the original scenario scheme. A scenario can be represented in a network by using the appropriate scenario scheme idiom as a building block and filling in the specifics of the scenario. The scenario scheme idiom thus provides the structure for modelling a scenario and it is furthermore annotated to enable the explaining of the network at a later stage.

The other three narrative idioms are the subscenario idiom, the variation idiom and the merged scenarios idiom. The subscenario idiom is very similar in structure to the scenario scheme idiom and it is intended to capture coherent substructures within a scenario, called subscenarios. The variation idiom can be used when a scenario includes small variations with no impact on the overall conclusion of a scenario; for instance, when it is uncertain how a burglar forced his way into a house, but the conclusion that this person stole items remains unchanged. The variation idiom provides a structure for modelling such variations without requiring an entirely new scenario structure for each variation, thus simplifying the complexity of the graph. Finally, the merged scenarios idiom can be applied to separate, competing scenario structures, to merge them into one Bayesian network structure for a case.

In Section 3.2, the quality of a scenario was investigated in a Bayesian network context. The quality of a scenario is expressed in terms of completeness, consistency and plausibility. The completeness of a scenario was defined relative to a scenario scheme idiom: when a scenario fits and completes a scenario scheme idiom, the scenario is complete. Consistency was defined in terms of probabilities: when a scenario contains two elements which can never occur simultaneously (the probability of the two both occurring is 0), then the scenario is inconsistent.

Finally, plausibility was also defined in terms of probabilities. The plausibility of an element of a scenario is equal to the probability of that element being true and the plausibility of the whole scenario is the probability of its scenario node. Plausibility can be represented in a network by calculating the appropriate numbers for the probability tables.

Finally, in Section 3.3, the construction method was proposed, using the idioms and the representation of scenario quality from the previous sections. The proposed method employed the notion that a scenario can be unfolded to the desired level of detail, thereby gradually constructing a Bayesian network with more detail. This process of unfolding is done by first representing an initial scenario with a scenario scheme idiom and then adding more detail by replacing elements of the scenario with subscenario structures. A structure representing each scenario is thus built up incrementally and finally the structures for alternative scenarios are merged to form one graph. Evidential nodes are then added to the structure and probabilities are elicited to reflect plausibility. This results in a fully specified Bayesian network which models the alternative scenarios, their quality and the evidence.

4. Case study: the Anjum case

In this chapter, a case study is presented as an evaluation of the method from Chapter 3. Using our method, we construct a Bayesian network for a case to test how well the method applies to a real case. In particular, we aim to evaluate with this case study the following list of criteria:

1. Does the combined method with scenarios and probabilities help to prevent tunnel vision?
2. Does the combined method with scenarios and probabilities help to prevent a good scenario being chosen over the true scenario?
3. Are appropriate scenario schemes available for the case at hand?
4. Does the five-step construction procedure help to simplify the construction of the Bayesian network?
5. Is the quality of a scenario adequately captured in the Bayesian network, using completeness, consistency and plausibility?

The first two criteria stem from a higher level goal that we had when developing this method: to combine scenarios and probabilities in such a way that advantages of each approach are put to their best use. A principal advantage of the scenario-based approach is that it helps to prevent tunnel vision by always requiring that several alternative scenarios are taken into account. With the first criterion, we aim to test whether this advantage is inherited by our combined method. On the other hand, the scenario-based approach comes with the common pitfall of selecting a good scenario over the true scenario. In a combined method, we hope to use the probabilistic approach such that this common pitfall can be avoided and the true scenario is selected, which is tested with the second criterion.

The third and fourth criteria are concerned with the construction of a Bayesian network, which is something that our method aims to simplify. In particular, the use of scenario scheme idioms should provide ready-made structures that can simplify the construction of the graph (criterion 3), while the five-step construction procedure aims to help a modeller keep an overview of the network while constructing it (criterion 4).

Finally, the fifth criterion is meant to test whether the quality of a scenario is captured in the Bayesian network. A scenario-based approach to reasoning with legal evidence compares alternative scenarios mainly on two properties: their relations to the evidence (or evidential support) and their quality. While a probabilistic approach lends itself well to capturing evidential support, representing the quality of a scenario in a Bayesian network is less straightforward and is something we specifically aimed to address in our method (see Section 3.2).

In the remainder of this chapter we first introduce the Anjum case (Section 4.1), which is then followed by five sections each discussing one step in the construction method for constructing a Bayesian network. These steps are representing the scenarios (Section 4.2), unfolding each scenario (Section 4.3), merging the structures (Section 4.4), including evidence nodes (Section 4.5) and specifying the probabilities (Section 4.6). After constructing the network, we take a brief look at some of the conclusions that can be drawn from the network (Section 4.7). This chapter concludes with a discussion, in which an evaluation of the method by means of the five criteria above is presented (Section 4.8).

4.1 The case

The case that will be modelled in this chapter is about the murder of Leo de Jager. This murder case is one of two murders that took place in the small Dutch town of Anjum, which were both investigated in the book by Crombag and Israëls (2008) as part of 'Project Gerede Twijfel' (Project Reasonable Doubt).[1] In this project, conducted by scholars from the VU University Amsterdam and the University of Maastricht, criminal cases in The Netherlands with a definitive conviction are investigated "if there is a real possibility that an innocent person was convicted". Our case study is based on this book as well as on the analysis by Bex (2011), which was in turn based on the book by Crombag and Israëls. Since these two books were our only source of information for this case, our results are undoubtedly influenced by the ideas presented by Crombag, Israëls and Bex. The aim of this case study is therefore not to evaluate the case objectively, but rather to evaluate our techniques by modelling this complex case in a Bayesian network.

On December 24th, 1997, Evert Beekman came into the police station to report a murder. Beekman said he had seen a dead body on the property of Marjan van der E. and that he had recognized the body as Leo de Jager. Furthermore, on Beekman's instructions, the police dug up the remains of another body in Marjan's garden, recognized as Herre Sturmans. In the book by Crombag and Israëls, the two murder cases are treated separately and separate scenarios are developed for each. It is the murder of Leo de Jager that will be discussed here.

Crombag and Israëls formulate four scenarios concerning the murder of Leo de Jager. In order to model a complete network for a case, the design method requires all relevant scenarios to be modelled. In this case study, the modelling is discussed

[1] www.projectgeredetwijfel.nl (in Dutch)

for two of the four scenarios, namely a scenario with Marjan as the perpetrator and a scenario with Marjan and Beekman working together. The other two scenarios concern the possibility that Beekman worked alone or that Leo's death was an accident, but these are by Crombag and Israëls considered less relevant alternatives. Throughout this chapter we use the fictitious names as used by Crombag and Israëls (2008) and Bex (2011). Only for the prime suspect, Marjan van der E., is her real name used.

4.1.1 The people involved

Marjan was the proprietor of a boarding house in Anjum. Leo rented a small house from Marjan in Moddergat and did some odd jobs around the boarding house. Beekman was a dealer in timber in Anjum and he knew Marjan because she had placed orders with him in the past. Other important persons involved in the case are Marga Waanders, who was staying in the boarding house at the time of the murder, and Eef Tasman, who did some administrative work for Marjan occasionally. Finally, Jaap Kuilstra had heard from Beekman about the murder, advised him to go to the police and came in with him to the station.

As it turned out later, Marjan, Beekman and Kuilstra had a cannabis operation in Marjan's barn. At the time that Beekman reported the murder, the police had found the operation a week before and Marjan was a suspect. However, she had denied any involvement and claimed that she had let the barn to someone else. She had then promised that she would show the police a contract of this agreement.

4.1.2 The evening of December 23rd

Initially, Beekman reported the events as follows. Marjan came to him on the evening of December 23rd to tell him that she had killed someone. This was around 7 in the evening. She returned to the boarding house and Beekman arrived there soon after. He talked to Waanders for a while, while Marjan was cleaning the hallway: she said that Leo had puked there and they weren't allowed to see. In all, Beekman and Waanders were talking for about 10 minutes. At some point, Waanders went into the hallway to get a washcloth and returned shortly after. Beekman later went into the hallway and saw Marjan scrubbing the floor. Beekman saw blood in the hallway and recognized a trail of blood which he thought might be from the back of a head as a body was dragged to the front door.

Then Marjan took Beekman to the front door, where he saw a dead body lying outside under tent canvas and recognized it as Leo's. Beekman reports that the victim's head was injured in 6 or 7 separate places, which he assumed was the result of hitting the head with a sharp object. Later, Beekman said that Marjan did not just tell him that she had killed someone, but that she had killed Leo. Furthermore, he admitted that he helped Marjan to wrap the body in the piece of canvas. He then returned to the boarding house at 2 a.m. to help Marjan drag the body to the front yard. Kuilstra confirms this story and explains that he had advised Beekman not to tell this part to the police.

When Marjan was first interrogated, she seemed too confused to say anything informative. In any case, throughout the investigative process she persistently claimed that she did not kill Leo, nor did she drug him. Marga Waanders, who was at first also a suspect, did give a statement right away. She said that Leo was at the boarding house when she arrived in the afternoon. She last saw him in the hallway around 6 p.m., talking to Marjan who was trying to convince him to stay in one of the rooms in the boarding house. In later interrogations, Waanders gave some more details about the events of December 23rd and 24th. In the afternoon, Leo seemed under the influence of something. Waanders also mentioned that she found him in the barn at some point and took him back to the house. Marjan then seemed agitated to find Leo back in the house. Marjan gave Leo a glass of warm water with jenever (Dutch gin), which she called 'a grog'. Later, when Marjan did not show up for dinner, Waanders took a look in the hallway and saw Marjan giving Leo another glass of jenever. Later, Waanders also said that she had 'images' of Marjan hitting Leo, but she said that these images do not mean that she actually saw this.

Waanders explicitly stated that she did not see any blood stains in the hallway. She did see Beekman when he came over, around dinner time. She then went to get a washcloth because her eyes were irritated. Later in the evening, Waanders saw two shadows standing outside, possibly Beekman and Marjan. She also saw someone digging a hole in the front yard some time on the 23rd or 24th of December. At the end of the evening, Marjan and Waanders had a drink together and went for a walk with their dogs, Waanders said to the police.

4.1.3 The evidence

Statements made by Beekman, Waanders, Marjan and Kuilstra serve as evidence in this case. Their main points were summarized in the description above. Additional information will be discussed as soon as it is of interest for the construction of the Bayesian network below. In this section, we present the key evidence other than the testimonies.

When investigating the boarding house on December 25th, the police found traces of blood in several places. Most of the blood traces were in the hallway. Furthermore, a wad with a bloody knot of hair was found in the trash can in Waanders' room. The police also found two hammers, a large one and a regular sized one, with watery bloodstains on them. These hammers were found in the barn. For each of these blood stains, a DNA match was found with Leo, though the profiles drawn from the analysed material were not complete. For the blood in the hallway the probability that this was from a random other person than Leo was estimated to be much less than 1 in a million. For the blood on a hammer, this probability was 1 in

100 for the hammer head and 1 in 1700 for the hammer handle (although the latter estimations have been disputed by another expert).

In the trash in the kitchen three empty strips of the medicine Temazepam were found and a strip with ten empty capsules that were cut open. Additional empty capsules were found in the trash, plus a medicine bottle in the name of Leo de Jager. An autopsy of Leo's body showed high concentrations of Temazepam and alcohol in Leo's blood. The level of Temazepam was far more than the amount advised for daily use and in fact far above the toxicity level. Temazepam is not lethal, but it will cause some strange behaviour. A pathologist concluded that the cause of death must have been a heavy blow to the head, leading to a fractured skull.

Finally, part of the evidence that was used in the trial concerning the Anjum case had to do with a bank fraud in which Marjan was supposedly involved. This may have been an additional motive for killing Leo, as Marjan seemed to be working towards transferring money away from Leo's account. To keep our model compact, we have chosen to leave out this motive and the evidence related to it.

4.2 Step 1: Representing each scenario

4.2.1 Scenario 1: Marjan killed Leo

The first scenario that is modelled here (the second scenario from Crombag and Israëls (2008)) is an elaborate version of what can be distilled from the police investigation. In this scenario, Marjan killed Leo and Beekman helped her move the body.

The initial first scenario goes as follows:

> Marjan had a cannabis operation. She used Leo as a front for this cannabis operation and then killed him. She then moved Leo's body with help from Beekman.

The first task is to select a scenario scheme idiom to which this scenario fits. Such a scheme is shown in Figure 4.1 (top): this scenario scheme describes a situation in which someone (X) is involved in criminal activities, but they use someone else (Y) as a front for these criminal activities. They then proceed to kill this person and move their body.

Using this scenario scheme idiom, the initial scenario can be modelled as in Figure 4.1 (bottom). This scenario is complete with respect to this scheme and it contains no inconsistencies. At a later stage, when probabilities are elicited, we will be able to capture the implausibility of Beekman being suddenly involved in moving Leo's body, while he is not involved in the rest of the scenario.

For each of the nodes in this structure, there is possibly a subscenario to unfold. This will be done in Section 4.3.1.

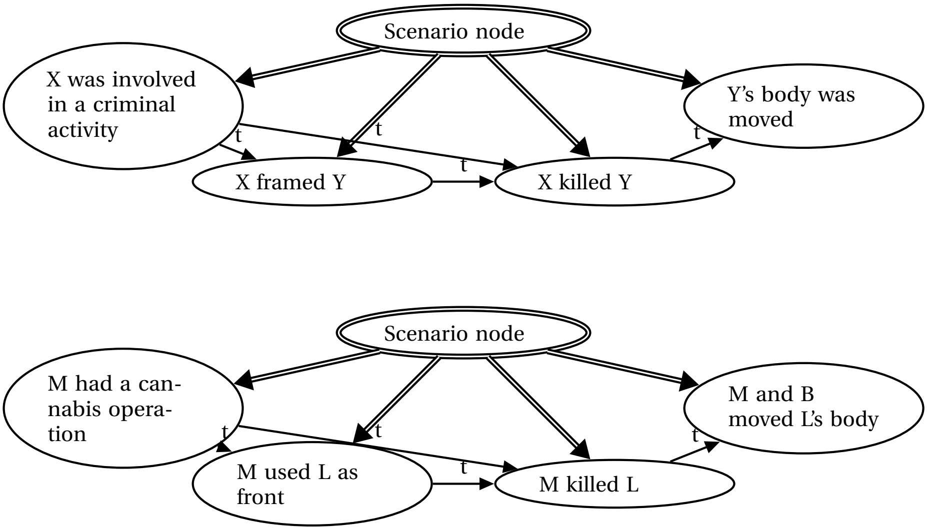


Figure 4.1: A scenario scheme idiom (top) and the initial scenario (bottom) for Marjan killing Leo.

4.2.2 Scenario 2: Beekman killed Leo

In an alternative scenario, Beekman killed Leo with Marjan's help. This is based on Crombag and Israël's third scenario. The initial second scenario goes as follows:

> Marjan and Beekman had a cannabis operation. Marjan used Leo as a front for this cannabis operation and Beekman then killed him. He then moved Leo's body to the yard.

This scenario fits the same scenario scheme idiom as the scheme shown in Figure 4.1 (top). The initial scenario can thus be modelled as shown in Figure 4.2. This scenario is complete with respect to this scheme and it contains no inconsistencies. Again, there are possibly subscenarios to unfold. This will be done in Section 4.3.3.

4.3 Step 2: Unfolding each scenario

To determine whether a node should be unfolded, the three questions from Section 3.3.1 are asked. In the sections that follow, the unfolding of specific nodes is discussed.

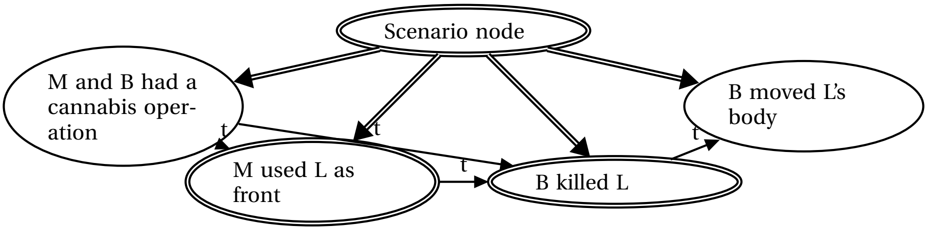


Figure 4.2: The initial scenario for Beekman killing Leo (made with the scenario scheme idiom from Figure 4.1 (top)).

4.3.1 Unfolding scenario 1

4.3.1.1 The cannabis operation

The leftmost node of the scenario about Marjan, `M had a cannabis operation`, has some evidence that can be connected to it directly (answering question 1 with yes): a police report from another investigation in which the police found a cannabis operation in her barn. Therefore, unfolding is not required. Of course it is still possible to unfold this node: the police report of the cannabis case surely relies on more detailed evidence about the cannabis operation. However, in order to keep the graph from getting too complex, those details are left out of this particular murder case.

4.3.2 Marjan used Leo as a front

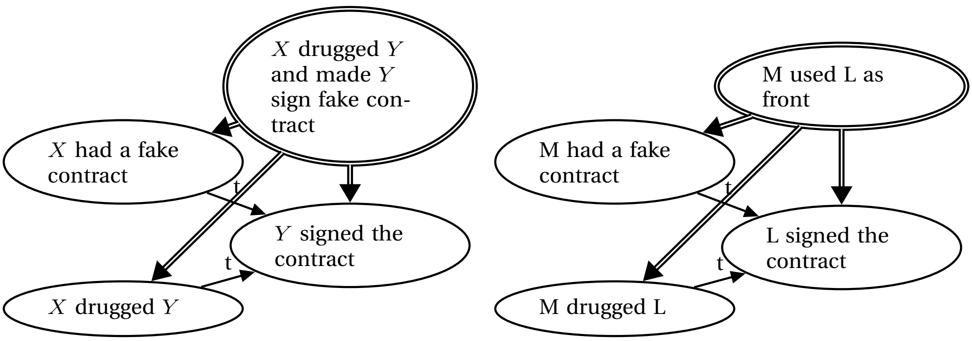


Figure 4.3: A scenario scheme idiom (left) and a subscenario (right) about Marjan using Leo as a front.

The node `M used L as front` has no evidence that can be connected to it directly (question 1: no). But there are some indications (question 2: yes) that Marjan was indeed planning to use Leo as a front for the cannabis operation: her accountant Tasman testified that he made up a false contract and this false contract was found in Marjan's house. By unfolding this node, the relation of this evidence with the state that L was to be a front for the cannabis operation can be specified.

Marjan needed to present a contract to the police to support her alibi that she had rented out her barn to someone else, so she had a false contract made. She drugged Leo, who then signed the contract. This can be modelled with a scenario scheme idiom as shown in Figure 4.3 (left), resulting in a subscenario structure as shown in Figure 4.3 (right). This subscenario is complete with respect to the scenario scheme idiom in Figure 4.3 (left) and it contains no inconsistencies.

The aforementioned evidence will be included as a final step in the construction method (after representing and merging the scenarios). For the nodes in this sub-

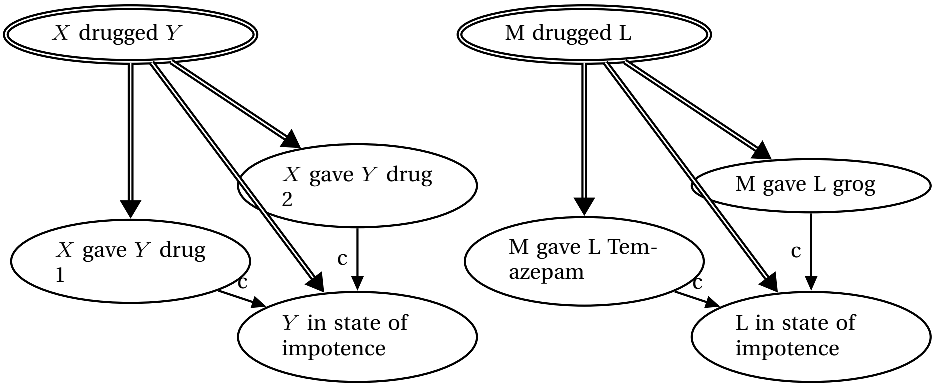


Figure 4.4: A scenario scheme idiom about drugging with two different drugs (left) and the subscenario for Marjan drugging Leo (right).

scenario, we can once again ask the three unfolding questions. For the nodes `M had a fake contract` and `L signed the contract`, there is direct evidence (question 1: yes), so they do not require unfolding. The node `M drugged L` certainly requires unfolding: there is no evidence for it directly (question 1: no) but there is evidence that Leo was drugged (Temazepam in his blood) and that Marjan had access to these drugs (bottles of Temazepam, question 2: yes).

A scenario scheme idiom for this subscenario is shown in Figure 4.4 (left), as well as the subscenario modelled with this scheme (right). This subscenario is complete with respect to this scheme and it contains no inconsistencies. The nodes in this subscenario require no further unfolding, since there is direct evidence for each of them.

4.3.2.1 Marjan killed Leo

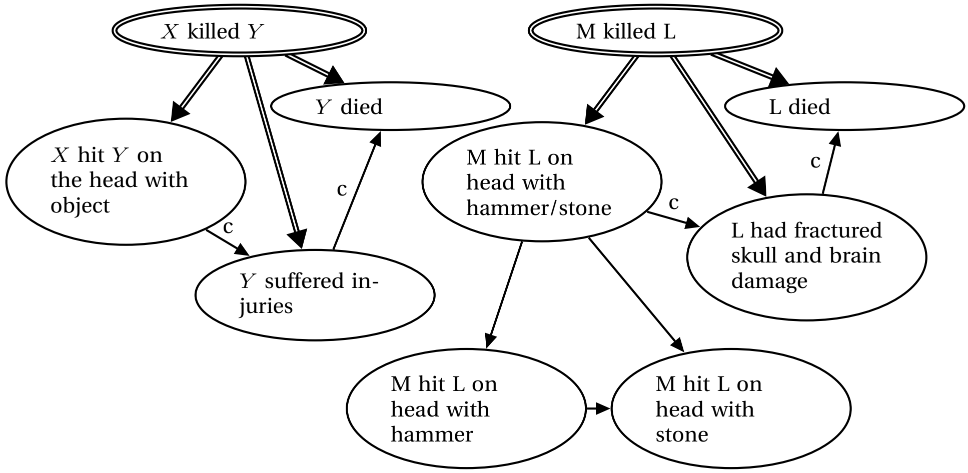


Figure 4.5: A scenario scheme idiom for someone hitting someone on the head (left) and the subscenario about Marjan hitting Leo on the head (right). The subscenario also includes a variation.

The node `M killed L` in the first scenario has no direct evidence (question 1: no) and more evidence can be included by unfolding it. In fact, several pieces of evidence can be included: there is blood on the hammer, a wad with blood and hair on it was found and there is an autopsy report about Leo's death (question 2: yes). Finally, there is Marjan's testimony, where she denies that she killed Leo.

Marjan supposedly killed Leo with a hit on the head with either a hammer or a stone, resulting in brain damage and a fractured skull, and ultimately Leo's death. This can be modelled with the scenario scheme idiom shown in Figure 4.5 (left), leading to a subscenario model as shown in Figure 4.5 (right). This subscenario also includes a variation, since it is unclear whether Marjan hit Leo with a hammer or with a stone (and there is evidence for both). In Figure 4.5, the variation idiom is used to represent Marjan hitting Leo with a hammer or a stone as mutually exclusive alternatives. The subscenario is complete with respect to the scenario scheme idiom that was used and it contains no inconsistencies. No further unfolding is required, since there is direct evidence for each of these nodes.

4.3.2.2 Marjan and Beekman moved Leo's body

Finally, in the first scenario there is the node `M and B moved L's body`. There is evidence that Marjan went to Beekman, but this is not direct evidence for the node

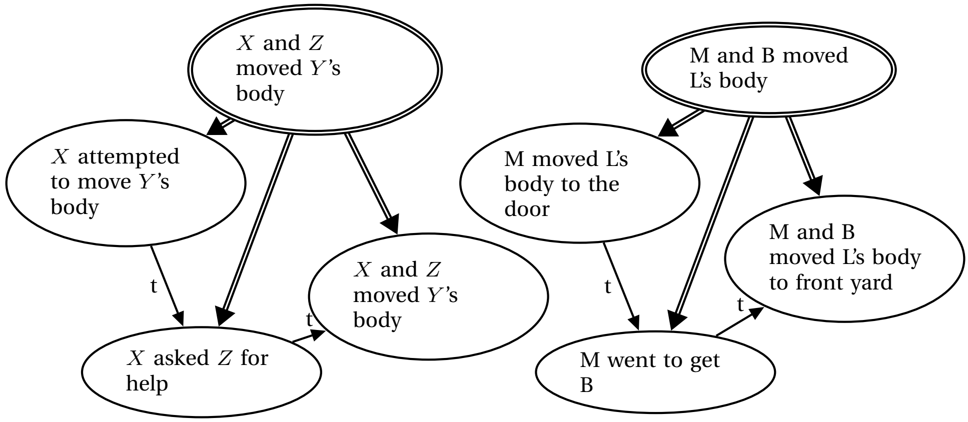


Figure 4.6: A scenario scheme idiom for someone asking for help with moving a body (left) and the subscenario about Marjan and Beekman moving the body (right).

(question 1: no) but only for part of the subscenario that can be unfolded here (question 2: yes). Furthermore, there is evidence that Leo's body was dragged to the front yard (Leo's body and a trail of his feet being dragged).

The subscenario about Marjan and Beekman moving the body is thought to be the following: Marjan first moved the body to the door (which is where Beekman says he saw it) and then went to Beekman for help, who then helped her move the body to the front yard. A scenario scheme idiom for this is shown in Figure 4.6 (left), in which someone attempts to move a body alone, goes to get help and then is helped in moving the body. This amounts to a subscenario structure as shown in Figure 4.6 (right). This subscenario is complete with respect to the scenario scheme idiom and it contains no inconsistencies. The plausibility of elements of this subscenario will be discussed in Section 4.6 below. No further unfolding is required.

4.3.2.3 The unfolded first scenario

The network structure for scenario 1 with the unfolded subscenarios is shown in Figure 4.7.

4.3.3 Unfolding the second scenario

4.3.3.1 Marjan and Beekman had a cannabis operation

The leftmost node of the initial second scenario in Figure 4.2 is `M and B had a cannabis operation`. Similar to the node `M had a cannabis operation` in the first

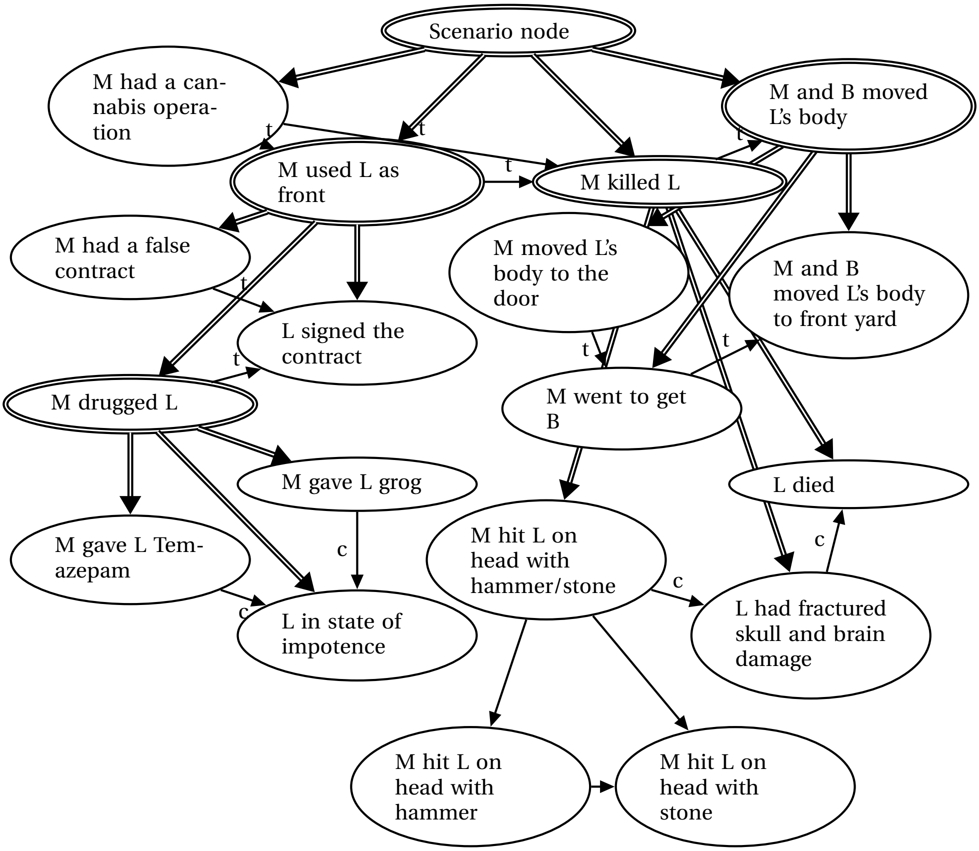


Figure 4.7: Scenario 1: Marjan killed Leo and Beekman helped move the body.

scenario, there is direct evidence for this node (question 1: yes) since the police was already aware of the cannabis operation in Marjan's barn. This node thus requires no unfolding.

4.3.3.2 Marjan used Leo as a front

This node also occurs in the first scenario and the subscenario that was unfolded there can be re-used in the second scenario.

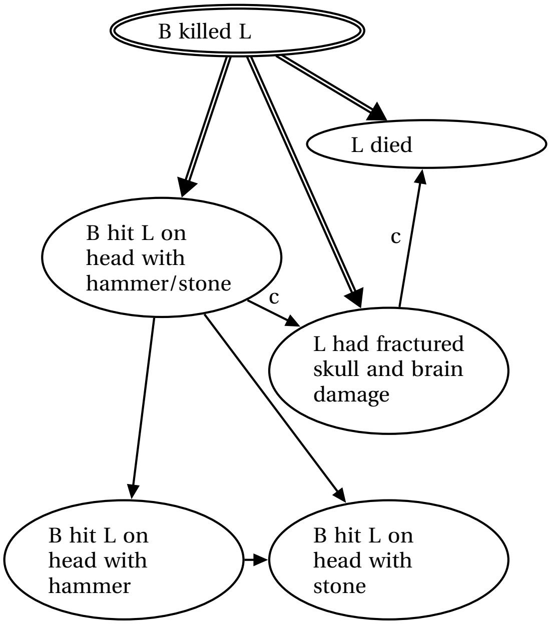


Figure 4.8: A subscenario about Beekman killing Leo.

4.3.3.3 Beekman killed Leo

The node `B killed L` in the second scenario can be unfolded in a similar fashion to the unfolding of `M killed L`, using the scenario scheme idiom from Figure 4.5 (left). This leads to a subscenario as shown in Figure 4.8. This subscenario is complete with respect to this scenario scheme idiom and has no inconsistencies.

4.3.3.4 Beekman moved Leo's body

Finally, the second scenario includes the node `B moved L's body`. There is direct evidence for this node in the form of Beekman's testimony (question 1: yes), so this node requires no unfolding.

4.3.3.5 The unfolded second scenario

The network structure for scenario 1 with the unfolded subscenarios is shown in Figure 4.9.

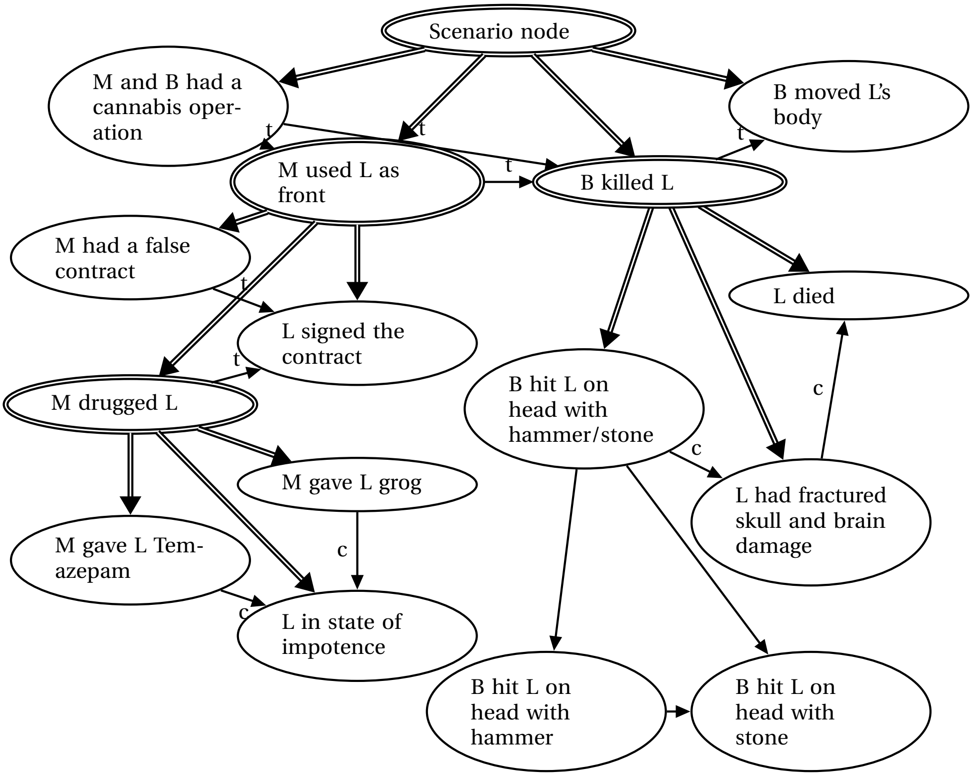


Figure 4.9: Scenario 2: Marjan drugged Leo and Beekman killed him.

4.4 Step 3: Merging the scenarios

For merging the two scenario structures in Figure 4.7 and 4.9, the merged scenarios idiom from Chapter 3.1 is used. The resulting structure is shown in Figure 4.10.

There are several overlapping elements in these scenarios, such as the node `M used L as front`, which occurs in both scenarios, and several elements of the subscenarios `M killed L` and `B killed L`. When a node occurs in both scenarios, the merged scenarios structure includes this node only once. For example, there is one node `M used L as front` in the merged structure and it is connected to both scenario nodes and to elements of each scenario according to the original connections in the two separate scenario structures. Similarly, the nodes `L died` and `L had fractured skull and brain damage` occur in the subscenario 'M killed L' as well as 'B killed L'. These nodes occur only once in the merged structure, with arrows drawn according to the original connections in the two subscenarios.

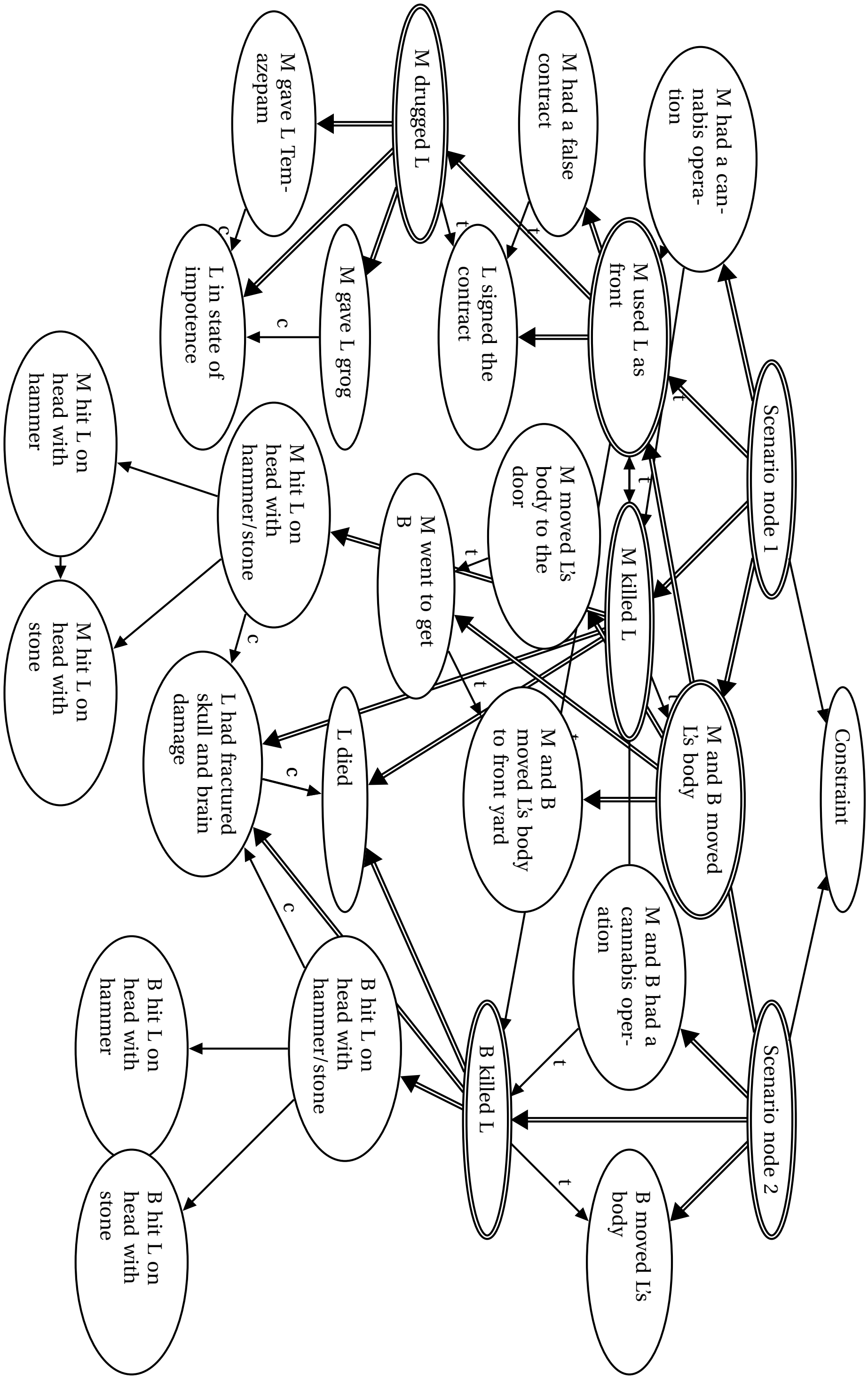


Figure 4.10: Scenario 1 and scenario 2 merged.

4.5 Step 4: Including evidence nodes

After merging the scenarios, evidential nodes are included. This comprises evidence about police findings, witness testimonies and forensic reports such as a toxicology report and an autopsy report. See Figure 4.11 for the resulting structure.

Each piece of evidence is modelled with the idioms by Fenton et al. (2013) This results in an evidential node accompanied by a node 'accuracy of evidence', abbreviated to `acc`. Each evidential node is connected to a node in the merged scenarios structure for which it is evidence. For example: the empty strips of Temazepam (evidence) are evidence for the hypothesis that Marjan gave Leo Temazepam (a node in the scenario). The accuracy-of-evidence node captures the possibility that a piece of evidence is not correct: there may have been a mistake in the lab (when it comes to DNA tests, for example) and a witness may lie.

Note that the accuracy of each witness is captured as a single node connected to all statements made by that witness: there is an accuracy node `acc` for Waanders, one for Beekman and one for Marjan. For both Beekman and Marjan the alibi evidence idiom by Fenton et al. (2013) was used to capture that their guilt in this case would influence the accuracy of their alibi evidence. According to this idiom, an arrow is drawn from the hypothesis about a witness' guilt (in our case, the corresponding scenario node) to the accuracy of this witness. As a result, if the scenario node for Marjan killing Leo holds, the accuracy of Marjan's testimony is low and if the scenario node for Beekman holds, the accuracy of Beekman's testimony is low.

4.6 Step 5: Specifying the probabilities

In this section we discuss the elicitation of probabilities for the network in Figure 4.11. Some, but not all, probabilities in the network will be discussed below to give an impression of the quantitative part of the Bayesian network, since the emphasis of our method is on the construction of the graph rather than the numbers.

Several numbers in the probability tables of the network are fixed because of the idioms that were used. For example, in the scenario scheme idiom and in the subscenario idiom, a number of connections are drawn as double arrows, signifying that a node is an element of the scenario or subscenario. For these nodes, part of the probability table is constrained. This means that, for example, for the leftmost node of the first scenario, `M had cannabis operation`, the probability table specifies that

$$\Pr(\texttt{M had cannabis operation = T}|\texttt{Scenario node 1 = T}) = 1.$$

Other probabilities remain to be determined. For example, for the node `M had cannabis operation`, what remains is $\Pr(\texttt{M had cannabis operation = T}|\texttt{ScN1 = F})$. These need to be elicited based on their plausibility (see Section 3.2). First,

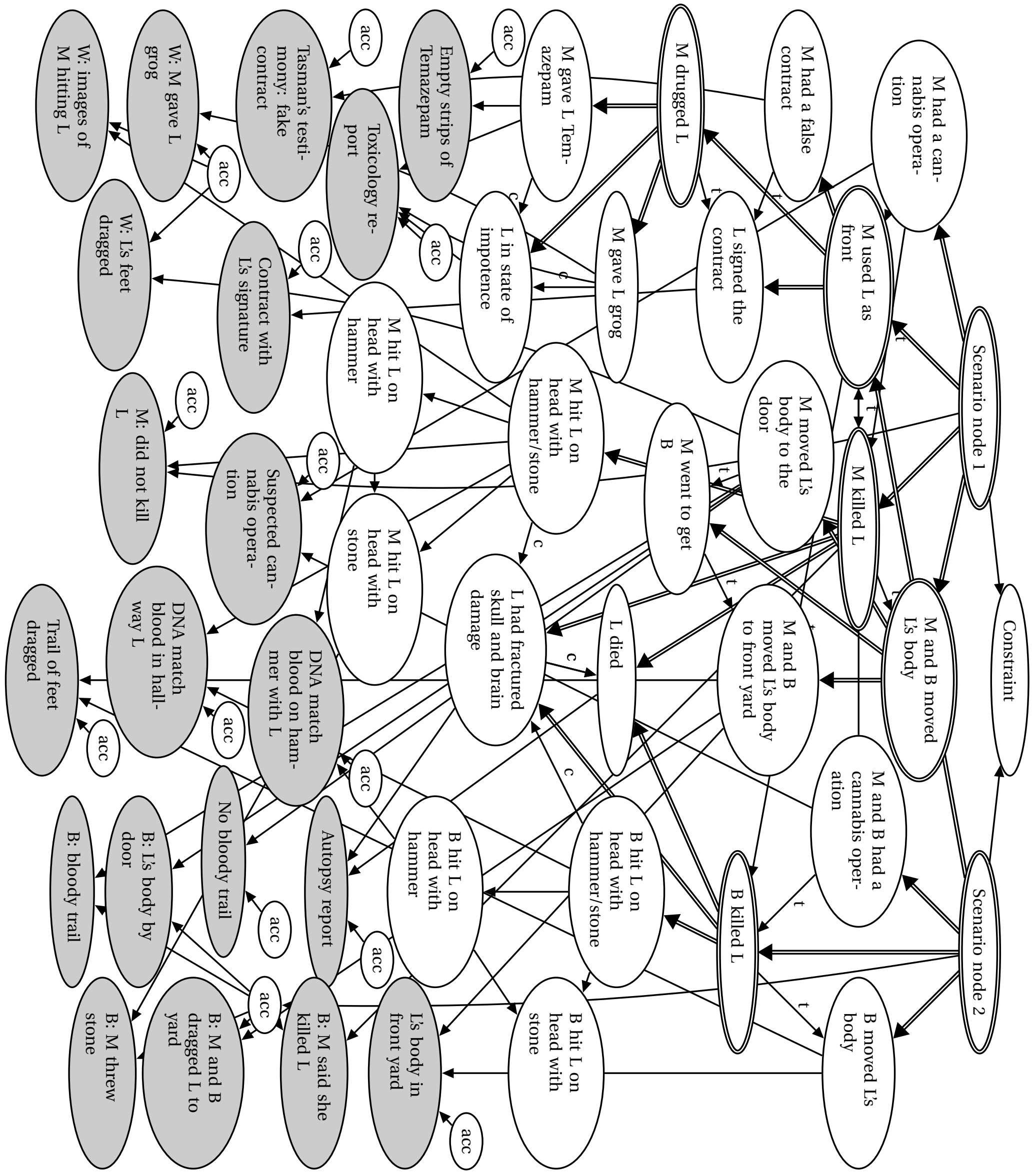


Figure 4.11: Scenario 1 and scenario 2 with evidence. Evidential nodes are indicated as grey nodes.

the plausibility of the element `M had cannabis operation` needs to be established. For example, one might estimate this to be an implausible element, with $\Pr(\texttt{M had cannabis operation = T}) = 0.001$. When the prior probability of the scenario node (the plausibility of the scenario as a whole) is also known, these numbers can be used to calculate the number needed for the probability table of `M had cannabis operation`:

$$
\begin{aligned}
&\Pr(\texttt{M had cannabis operation = T}|\texttt{ScN1 = F})\\
&= \frac{\Pr(\texttt{M had cannabis operation}) - \Pr(\texttt{ScN1 = T})}{\Pr(\texttt{ScN1 = F})}\\
&= \frac{0.001 - \Pr(\texttt{ScN1 = T})}{\Pr(\texttt{ScN1 = F})}\\
&= \frac{0.001 - \Pr(\texttt{ScN1 = T})}{1 - \Pr(\texttt{ScN1 = T})}.
\end{aligned}
$$

Suppose the plausibility of the scenario as a whole is 0.0001, then the required probability for `M had a cannabis operation` will be as follows:

$$\Pr(\texttt{M had cannabis operation = T}|\texttt{ScN1 = F}) = \frac{0.001 - 0.0001}{1 - 0.0001} = 0.0009.$$

Similarly, the plausibility of other elements in the scenario can be used to calculate the numbers required for their probability tables. This way, it can also be modelled that the element `M and B moved L's body` is implausible given the element `M killed L`. To represent this, we want to model it such that the plausibility of `M and B moved L's body` given `M killed L` is 0.001, that is, $\Pr(\texttt{M and B moved L's body = T}|\ \texttt{M killed L = T}) = 0.001$. Again, the required number for the probability table of `M and B moved L's body` can be calculated once the plausibility of the scenario and of `M killed L` are known:

$$
\begin{aligned}
&\Pr(\texttt{M and B moved L's body = T}|\texttt{M killed L = T, ScN = F}) =\\
&\Pr(\texttt{M and B moved L's body = T}|\texttt{M killed L = T})\cdot\\
&\frac{\Pr(\texttt{M killed L= T}) - \Pr(\texttt{ScN = T})}{\Pr(\texttt{M killed L = T}|\texttt{ScN = F}) \cdot \Pr(\texttt{ScN = F})}\\
&= 0.001 \cdot \frac{\Pr(\texttt{M killed L= T}) - \Pr(\texttt{ScN = T})}{\Pr(\texttt{M killed L = T}|\texttt{ScN = F}) \cdot (1 - \Pr(\texttt{ScN = T}))}.
\end{aligned}
$$

Sometimes the plausibility of elements is available within the evidence. For example, consider the node `L was in a state of impotence` in the subscenario `M drugged`

`L`. In the toxicology report, an expert stated that given that Leo had alcohol and high amounts of Temazepam in his blood, he would be in a state of impotence. This is a qualification of $P($`L was in state of impotence = T`$|$ `M gave L grog`, `M gave L Temazepam = T`$)$, which can be translated to a number using a verbal scale (Renooij, 2001). Then, from this plausibility, the required number for the probability table of `L was in a state of impotence` can again be calculated as explained above.

Similarly, the probabilities within the subscenario about Leo's death follow for a large part from the autopsy report, such as for the node `L died`, connected to `L had a fractured skull and brain damage`.

As for the accuracy of evidence, forensic evidence is often accompanied with numbers reporting the accuracy of the lab. The accuracy of other evidential data may be more difficult to estimate. For example, Waanders' testimonies changed a lot during the investigation, so they have been estimated to be less reliable.

Finally, after filling in all probabilities underlying the network, a sensitivity analysis (Jensen and Nielsen, 2007) can be performed on nodes with disputable probability tables. Such a sensitivity analysis determines the influence of a change in probabilities given that all other probabilities remain fixed. When multiple nodes in a network have disputable probabilities, a sensitivity analysis on multiple nodes should take into account any configuration of changed probabilities.

4.7 Conclusions drawn from the network

With the model that was constructed in the previous sections, conclusions about the case can be drawn. This also serves as feedback on whether the case has been modelled as intended. To perform calculations in the network, our model has been specified in the Bayesian network tool GeNIe 2.0[2]. Our model is available for download at www.charlottevlek.nl/networks. Running our network in GeNIe 2.0 shows that, given all evidence, the scenario that Beekman killed Leo is more likely than Marjan killing Leo.

By gradually entering the evidence one by one, the network gives insight into the case, explicating how this conclusion was reached. Without any evidence in the network, both scenario nodes (Marjan killed Leo and Beekman killed Leo) are very unlikely, as their prior probability was estimated to be 0.0001. For the purpose of readability, the remainder of the numbers reported in this section are rounded to four decimals. Loosely following the police investigation, we start by entering Beekman's testimonies. This leads to a somewhat higher probability for the scenario with Marjan as a killer (0.0016), while the scenario about Beekman is at a probability of 0.0005. Next Marjan's testimony is inserted in the network, but the numbers do not change: since her accuracy is connected to the scenario node, her testimony is taken to be most probably (0.9873) inaccurate.

[2] GeNIe 2.0 is available for free on www.bayesfusion.com

By adding more evidence to support (parts) of the scenario with Marjan as the killer, this scenario node obtains a higher probability. With the evidence of Marjan being a suspect for the cannabis operation, the probability of scenario node 1 becomes 0.0125 (scenario about Beekman: now 0.0036). Adding evidence about a contract with Leo's signature being found and Tasman's testimony about the fake contract, the probability of Marjan being the killer becomes 0.2718 (versus Beekman: now 0.0788). Various pieces of evidence were found to support that Marjan drugged Leo: empty strips of Temazepam, a toxicology report and Waanders' testimony that Marjan gave Leo a grog. This results in a very high probability (0.9807) for the subscenario node about Marjan drugging Leo and a high probability (0.9204) for the subscenario about Marjan using Leo as a front. Both scenario nodes now become more likely: the scenario node for Marjan killing Leo now has probability 0.3962 and the scenario node for Beekman has probability 0.1149.

Next, a number of forensic pieces of evidence about the killing itself are entered: a DNA match with Leo and the blood that was found on the hammer, a DNA match with Leo and the blood in the hallway, Leo's dead body in the yard and a trail of his feet being dragged through the yard, and the autopsy report. Both scenarios again become more probable, with the probability of scenario node 1 on 0.5792 and the scenario node for Beekman on 0.4182.

Note that at this point no evidence has been included about *who* killed Leo, except for a testimony from Beekman, stating that Marjan said she killed Leo. The evidence that was entered so far all supports both scenarios. A number of relevant statements were made by Waanders: she had 'images' of Marjan hitting Leo, and she saw Leo's feet being dragged away through the hallway. Entering these testimonies leads to a much higher probability for the scenario in which Marjan did it: 0.8991, while the scenario about Beekman is now on a probability of 0.0990.

Finally, a crucial piece of evidence turns out to be the fact that there was no bloody trail found in the hallway, even though Beekman said in his testimony that he saw such a trail. Taking into account that a scenario in which Marjan killed Leo would lead to such a bloody trail (she supposedly dragged Leo outside the door by herself) and that other blood traces were very obvious and easy to find for the forensic team, this piece of evidence contradicts the scenario about Marjan and it furthermore compromises Beekman's accuracy. The probability of Beekman's accuracy having value true is now 0.1470 instead of 0.9007 and the probability of Waanders's accuracy having value true is now 0.7654 instead of 0.9686. The updated probabilities of the scenario nodes are now 0.1467 for the scenario about Marjan and 0.8500 for the scenario about Beekman.

The network thus shows that there is not much evidence to really distinguish between the two scenarios, although the scenario for Marjan being the killer is the more likely one for large subsets of the evidence. As soon as evidence is entered to disprove some of Beekman's testimonies, the probabilities change to show that Beekman is much more likely to be the killer, with some room for doubt. Though this model clearly includes many subjective probabilities, it reflects the conclusions by Crombag and Israëls (2008).

4.8 Discussion of the case study

The case study presented in this chapter was meant as an evaluation of our method from Chapter 3. In this discussion, we present our findings based on this case study. In the following subsections, each of the criteria from the introduction of this chapter is evaluated.

4.8.1 Criterion 1: Tunnel vision

The first criterion is whether the method helps to prevent tunnel vision. In the investigation of the Anjum murder case, one scenario, that of Marjan killing Leo, dominated the case. There is thus a clear risk of tunnel vision when the focus lies too much on only this scenario. In our case study, an alternative scenario about Beekman killing Leo was modelled since the method always asks for more than one scenario to be represented. This reduces the risk of tunnel vision.

There is room for improvement in the prevention of tunnel vision, by making our approach more dynamic in that it could better handle changing and expanding scenarios during the investigative process. Currently, our method assumes that in an earlier phase, a number of scenarios were formulated. The method then results in a Bayesian network modelling these scenarios. However, during an investigation, scenarios about what happened are gradually formed. The same happens for the Anjum case in Crombag and Israëls (2008), for example concerning the killing of Leo. For a long time it remains unclear at what time Leo was killed and by whom. Based on various testimonies by Beekman and Waanders, multiple possibilities are taken into account, but only two possibilities deserve further investigation: either Leo was killed by Marjan some time before Beekman came to her house on the evening of December 23rd, or Leo was killed by Beekman some time later that evening. These two possibilities are the core of the two scenarios modelled in Section 4.2, while the more unlikely possibilities never made it into the network. Modelling this process of gradually constructing a theory during the investigation is a topic for future work. Moreover, an extended method that can help to find multiple scenarios during the process of investigation could actively help to prevent tunnel vision.

4.8.2 Criterion 2: A good scenario versus the true scenario

The second criterion is whether the method helps to prevent a good scenario being chosen over the true scenario. In the actual trial for the Anjum case, Marjan was convicted for the murder. Only in the thorough investigation by Crombag and Israëls (2008) did the scenario about Beekman turn up as a somewhat better explanation of the evidence. This conclusion is reflected in our network, as discussed in Section 4.7. It seems that the scenario in which Marjan killed Leo was good enough to be believed by many people for a long time: it might be that this was the good scenario that pushed out the true scenario in the trial.

A closer look at the network in Section 4.7 revealed how Marjan being the killer came to be so believable: most of the evidence supported (or at least, did not disprove) this scenario. However, as soon as some evidence was instantiated in the network to disprove one of Beekman's testimonies, the probabilities flipped, showing that Beekman probably lied and most likely committed the murder himself. Whereas a judge or jury might find it difficult to let go of their earlier belief that Marjan was the killer, the network reveals that Beekman's testimonies are likely to be false given all evidence. Formalizing the scenarios in a probabilistic framework can thus help to choose the more likely scenario over a good scenario.

There is one clear limitation to this formalization: it requires a high level of precision, both for the required numbers and for the structure of the graph. There are elicitation techniques to help find all the numbers (see Section 2.2), but it is not clear how successful these techniques are for this particular application to scenarios in the legal field. We argue that a Bayesian network is a useful tool to structure a subjective thought process and to compare relative probabilities of multiple scenarios. The risk of such a formalization is the illusion of objectivity it creates: an elaborate Bayesian network for a case might appear impressive to someone who was not involved in the construction. Quantifying all probabilities with numbers may lead to the impression that all these numbers are really known: unfortunately, in a Bayesian network there is no way to indicate that a probability is unknown. Our method is explicitly not intended to calculate any objective posterior probability, but rather to help formalize the subjective considerations of a judge or jury. Such a formalisation can help to find errors in reasoning such as a good scenario pushing out a true scenario.

4.8.3 Criterion 3: scenario scheme idioms

The third criterion is whether appropriate scenario schemes were available for the case at hand. In the Anjum murder case, scenario scheme idioms were used to model each scenario and its subscenarios. For example, we used a scenario scheme idiom about someone framing another person for their own criminal actions and then killing this person. In addition, scenario scheme idioms about framing, drugging and killing were used to model subscenarios. These scenario scheme idioms were developed 'on the fly', specifically with the application to this case in mind. As a result, these idioms are rather specific, whereas ideally a database with generic scenario idioms would be available to be used on various cases. Despite this issue of specificity, we found that the use of scenario scheme idioms systematises the construction of the network, even when they are developed on the fly, since the modeller is forced to take a step back and consider the underlying pattern of the scenario. Moreover, compared to a previous version of this case study (Vlek et al., 2014a) in which the method did not yet include the use of scenario schemes, the elicitation of the graphical structure without scenario scheme idioms was more ad hoc and subject to changes as the construction process progressed. We therefore conclude that scenario scheme idioms were available and work well in practice, but could be improved and would be easier to work with if a database of scenario scheme idioms were available.

The development of scenario scheme idioms on the fly could be further improved when more directions for their graphical structure are provided. Since the connections in a scenario scheme are often implicit, a modeller needs to think about the connections between nodes when specifying a scenario scheme idiom. This is of course no longer an issue when a database of scenario scheme idioms is available.

4.8.4 Criterion 4: the construction procedure

The fourth criterion is whether the five-step construction procedure helps to simplify the construction of the Bayesian network. During the construction of the network for the Anjum case, the five-step construction procedure helped to keep track of the modelling process. In particular, the resulting scenario structures in Figures 4.7 and 4.9 are quite complex, but the modelling process proceeded in an orderly fashion since these network structures could be gradually constructed by unfolding each scenario in the Anjum case. The process of unfolding thus helps to keep an overview of the case, even when constructing complex networks.

The construction procedure furthermore structures the construction process by first modelling the scenarios, then adding evidence and finally specifying the probabilities. The latter step is perhaps unreasonably large compared to the other steps: specifying all probabilities is a lot of work. While the focus of our method is currently not on the probabilities but rather on the graphical structure, the method could benefit from splitting up the fifth step of eliciting the probabilities into more concrete directions as to how to elicit these numbers.

4.8.5 Criterion 5: scenario quality

Finally, the fifth criterion is whether the quality of a scenario is adequately captured in the Bayesian network. With our method, the quality of a scenario could be captured in a network in terms of completeness, consistency and plausibility. In the Anjum case, both scenarios and their subscenarios were found to be complete, though this may have to do with our developing scenario scheme idioms on the fly, specifically tailored to fit the scenario. Both scenarios were also consistent and their completeness and consistency could be adequately represented as such in the

Bayesian network. The first scenario contained the implausible connection from Marjan killing Leo to Beekman then helping her to move the body: why would Beekman do this if he were not otherwise involved in the murder? This implausibility could be captured in the network by setting the numbers such that they result in a low probability for $\Pr(\texttt{M and B moved L's body} | \texttt{M killed L})$. Whereas our previous version of this case study (Vlek et al., 2014a) did not specify this connection as implausible per se, its implausibility could now clearly be captured in the network. However, capturing the plausibility of each scenario as a whole was not straightforward, mostly because it was difficult to assess the plausibility of each scenario in the first place. The quality of a scenario in terms of completeness, consistency and plausibility is thus adequately represented in the network, but more guidance on assessing the plausibility of a scenario would be helpful.

When specifying the probability tables of elements based on their plausibility, this process requires numerous calculations. Moreover, in the Anjum case we found that these calculations often have little effect due to low prior probabilities for the scenario nodes, meaning that the outcome is a number similar to the plausibility that went into the calculation. Our method could be greatly improved when these calculations no longer need to be done manually, but instead the plausibility of scenario elements could just be directly entered into a network. This could possibly be realized by developing a software tool that builds on Bayesian networks but allows for the input of plausibility rather than probabilities, performing some additional calculations in the background to find the required probabilities for the Bayesian network.

Part III: Explaining Bayesian networks with scenarios

5. Extending the construction method to explain Bayesian networks with scenarios

When capturing an entire case with several alternative scenarios in a Bayesian network, the result is often a large and complex structure. See, for example, the results from the Anjum case study in Chapter 4 of this thesis. Presenting such a network directly to a judge or jury is not an option, due to its complexity. And whereas the resulting networks work well to calculate the posterior probability of each scenario, only presenting the outcome of such calculations to a judge or jury is not appropriate, since this would essentially put the assessment of the case in the hands of the modeller. Some explanation of a network is thus needed.

Fenton and Neil (2011) compare the use of Bayesian networks in court to the use of a calculator which can help perform complex calculations. They argue that to understand the result produced by a calculator, it is not necessary to understand the exact workings of the calculator itself. Instead, one should understand the input that went into the calculator, since this is what determines the outcome. Therefore, what is needed to understand the results of a Bayesian network, according to Fenton and Neil (2011), is an understanding of the relations between variables as they were modelled in the network, and their probabilities.

In this chapter, we propose to extend our previous method with techniques for explaining what has been modelled in the network in terms of scenarios. According to Lacave and Díez (2002), three types of content in a Bayesian network can and should be explained: the model, the evidence and the reasoning. By explaining which scenarios are in the network, we aim to explain the content of the model and by explaining how each piece of evidence supports each scenario, we aim to explain the evidence and the reasoning from the evidence to the conclusion.

Scenarios have been found to provide a way to make sense of the evidence for a judge or jury (Pennington and Hastie, 1993). By communicating which scenarios are in the network and how they are represented in relation to the evidence, we aim to give a coherent overview of the contents of the Bayesian network. Moreover, by communicating the ingredients that are needed for a scenario-based assessment of the case, a judge or jury should be able to draw their own conclusions about the case in terms of a scenarios.

Our goal in this chapter is to extract from a Bayesian network the ingredients that are needed for a scenario-based approach, namely: (1) the alternative scenarios in the case, (2) their quality and (3) their evidential support. By enhancing these three elements with probabilistic information from the Bayesian network, a judge or jury can consider the case in terms of scenarios, while taking the probabilistic information into account as well.

In what follows, we propose methods for obtaining from the network the three ingredients described above. In Section 5.1 we propose a method for extracting a scenario in text-form from a Bayesian network. In Section 5.2 we discuss how evidential support for each scenario can be reported. In Section 5.3 we consider the reporting of scenario quality as it was modelled in the network. Finally we propose a reporting format in Section 5.4, which can be used to report these three elements.

5.1 Extracting scenarios

In this section, a technique is proposed for extracting several alternative scenarios in text-form from a Bayesian network. In Section 5.4, these texts are used to form a report about the contents of the network, enhanced with (probabilistic) information about scenario quality and evidential support. By presenting the alternative scenarios that have been modelled in a network, the contents of that network can be communicated to a judge or juror.

To extract scenarios from a network, we employ properties the network has as a result of the construction method from Chapter 3, which is why the explanation techniques in this chapter are considered an extension of the method from Chapter 3.

Below, we first discuss how the network structure can be used to identify alternative scenarios, subscenarios, variations and evidence in the network (Section 5.1.1). We then show how these structures can be translated to text in Section 5.1.2.

5.1.1 Identifying structures in the network

A network that is constructed with the method from Chapter 3 contains special nodes that can be used to identify various structures in the network: the scenario node, the subscenario node, the disjunction node, the constraint node and evidence nodes. These can be used as follows to identify scenarios, subscenarios, variations and evidence in the network:

- Each scenario node represents a scenario. The children of that scenario node minus the constraint node form the elements of the scenario. These elements may include subscenario nodes or disjunction nodes;

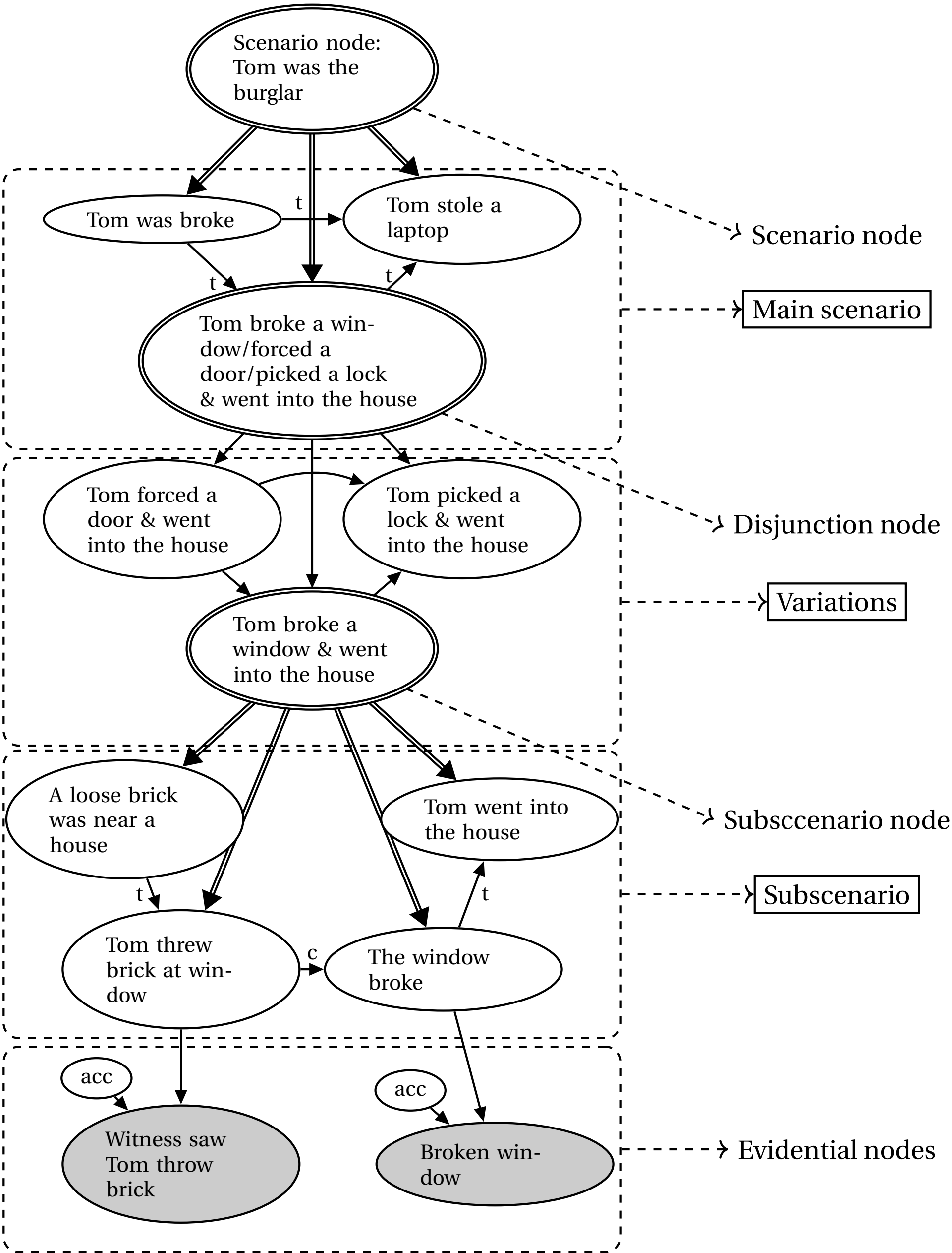


Figure 5.1: Identifying various structures in a network.

- Each subscenario node represents a subscenario. The children of the subscenario node (including possibly further subscenario nodes or disjunction nodes) form the elements of the subscenario;
- A disjunction node represents a variation. The children of the disjunction node (including possibly subscenario nodes or further disjunction nodes) form the variations.
- Evidence nodes represent pieces of evidence;

An illustration of these identifications of structures is shown in Figure 5.1, which models a burglary scenario from Section 3, now with some evidential nodes as an example. In this illustration, we see the main scenario at the top, identified by (but not including) a scenario node. This main scenario includes a disjunction node as one of its elements, from which three variations can be identified in the structure. Among these variations is a subscenario node, which identifies a subscenario about Tom breaking a window. Finally, there are evidential nodes at the bottom.

5.1.2 Translating scenario structures to text

As described in the previous subsection, scenarios, subscenarios and variations can be identified in a Bayesian network. In this section, we will first focus on how the network structure of a scenario can be translated to text, illustrated with some example structures in Figure 5.2. This is followed by a treatment of subscenarios and variations. Again we use properties of the network that are specific to our construction method from Section 3, namely:

1. Each node that is an element of a (sub)scenario represents a proposition. This includes subscenario nodes and disjunction nodes. Variation nodes also represent propositions. To retrieve a proposition from a node, the name of that node can be used directly as a proposition;
2. Arrows between elements of a (sub)scenario correspond to connections in the (sub)scenario;
3. Arrows between elements of a (sub)scenario are labelled to indicate the type of connection; in this chapter we assume that a connection can be either causal ('c') or temporal ('t').

Nodes to propositions Consider the example in Figure 5.2(a). Each node represents a proposition, given by the name of the variable in the node: P_i. Even though technically, the value-assignment $P_i = \mathtt{T}$ represents the proposition, the name of the variable can be used to compose a text.

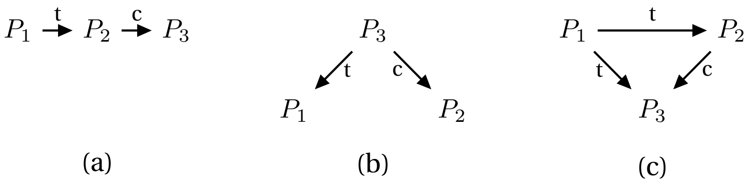


Figure 5.2: Some example structures.

Example (Variable names as propositions). Consider the node `Tom was broke` in Figure 5.1. The name of this variable can be used directly in a text as the proposition 'Tom was broke'.

Labels to connectives Since the arrows between nodes represent connections in the scenario, they can be used to order and connect propositions in the text. The ordering of propositions is discussed in more depth below; to connect propositions we propose to use connectives 'therefore' for 'c'-labels and 'then' for 't'-labels. A text is then composed by placing the propositions in order, with the appropriate connectives in between.

Example (Connectives). The example structure in Figure 5.2(a) is translated to the following text: 'P_1. Then P_2. Therefore P_3.'

Multiple parents or children In Figure 5.2(b), a structure is shown in which a node has multiple children. When this happens, propositions that are all children of the same node can be grouped together using 'and'. When arrows connecting various children to one parent have different labels, we propose to use the 'c'-label for our translation to text, since we consider the causal connection to be of more relevance for the scenario.

Example (Multiple children). The example structure in Figure 5.2(b) is translated to the following text: 'P_3. Therefore, P_1 and P_2.'

When a node has multiple parents, they can be grouped in a similar fashion. However, whereas multiple children can always be grouped with 'and', for multiple parents situations may occur in which it is more appropriate to use 'or'. This depends on the probability table of the child node and whether the 'explaining away' effect occurs (Wellman and Henrion, 1993). This is when the two parents A and B (now interpreted as causes of child C) form alternative explanations of C. Consider for example A: Tom threw a stone at the window and C: the window broke, but now with the addition of an alternative cause B: a bird flew into the window. In this case, knowing A (Tom threw a brick at the window) reduces the need to assume B (a bird flew into the window) to explain the effect C (the window broke). This effect of explaining away occurs when the effect on C of knowing B to be true is smaller when it is already known that A is true, that is, when the following holds (Wellman and Henrion, 1993):

$$\frac{\Pr(C = \mathrm{T}|A = \mathrm{T}, B = \mathrm{T})}{\Pr(C = \mathrm{T}|A = \mathrm{T}, B = \mathrm{F})} \leq \frac{\Pr(C = \mathrm{T}|A = \mathrm{F}, B = \mathrm{T})}{\Pr(C = \mathrm{T}|A = \mathrm{F}, B = \mathrm{F})}$$

When this is the case, alternatives A and B are grouped with 'or'.

When multiple parents are related In Figure 5.2(c), an example is shown where multiple parents of a node are also related (they are said to married): P_1 and P_2 are parents of P_3, but there is also a direct connection between P_1 and P_2. An appropriate translation of this is not straightforward, since it is not clear whether the parents of P_3 should be grouped, or the children of P_1.

Instead of grouping either parents or children, we propose to favour the longer path that exists between P_1 and P_3 via P_2, ignoring the direct connection between P_1 and P_3. The intuition behind this is as follows: though the connection between P_1 and P_3 is needed to represent that there is an influence even when P_2 is instantiated, when P_2 remains unobserved this influence is already contained in the longer path. Note that since P_2 is an element of a (sub)scenario, it will never be instantiated.

Example (Related parents). The example structure in Figure 5.2(c) is translated to the following text: 'P_1. Then, P_2. Therefore, P_3.'

In general, for a node that has multiple parents which also have a directed path between them, we propose to use the longest path from a parent to the child node and ignore any shorter connections there may be.

Producing text for any structure Consider a Bayesian network graph $(\mathcal{V}, \mathcal{E})$, and let $\mathcal{P} \subseteq \mathcal{V}$ denote the set of propositions in a (sub)scenario as identified by a (sub)scenario node (a subset of the set of all nodes in the Bayesian network). Let $(\mathcal{P}, \mathcal{C})$ be the full subgraph for the set of propositions $\mathcal{P}$. As an illustration, the subgraph for the main scenario structure from the network in Figure 5.1 is shown in Figure 5.3. $\mathcal{P}$ is thus the set of nodes that represent the elements of the scenario and $\mathcal{C}$ is the set of connections between these nodes (a subset of all connections in the Bayesian network). The subgraph $(\mathcal{P}, \mathcal{C})$ is the structure that needs to be translated to text. Using the ideas described above, we now propose a general procedure for translating any such structure $(\mathcal{P}, \mathcal{C})$ to text:

1. Let Q_1 be the set of nodes in $\mathcal{P}$ that have no parents in $\mathcal{P}$:

$$Q_1 = \{P_i \in \mathcal{P} | \neg \exists P_j \in \mathcal{P} : (P_j, P_i) \in \mathcal{C}\}.$$

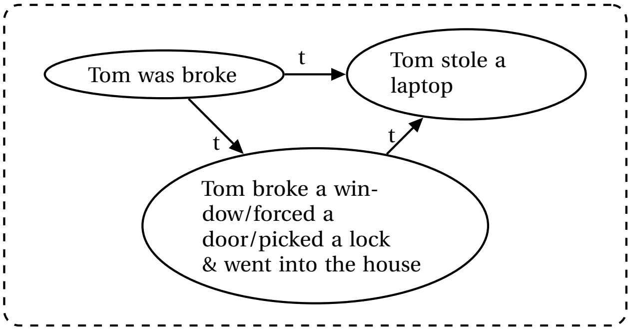


Figure 5.3: The scenario structure, isolated from the network in Figure 5.1.

Conjoin all propositions in Q_1 using 'and' or 'or', depending on whether the explaining away effect occurs.

2. let Q'_{i+1} be the set of children of nodes in Q_i:

$$Q'_{i+1} = \{P_i \in \mathcal{P} | \exists P_j \in Q_i : (P_j, P_i) \in \mathcal{C}\}.$$

Define a new set Q_{i+1} which is found by removing any node P_i from Q'_{i+1} when there is a directed path from a node in Q_i to P_i via another node in Q'_{i+1}:

$$Q_{i+1} = \{P_i \in Q'_{i+1} | \neg \exists P_j \in Q_{i+1} : (P_j, P_i) \in \mathcal{C}\}.$$

Conjoin all propositions in Q_{i+1} using 'and' or 'or' and place this string of text after the previous text, using connective 'therefore' when there is at least one 'c'-connection between a node in Q_i and a node in Q_{i+1}, and connective 'then' otherwise.

3. Stop when Q_{i+1} is an empty set.

This process will always stop since a Bayesian network is a directed acyclic graph. This procedure does the following: in step one, all elements of the scenario are grouped together that have no parents within the scenario. In step two, children of this first group of elements are grouped together, but any node that is connected via another child in this group is removed. This second step is repeated until no further child nodes are left.

Example (Translating a scenario to text). Applied to the scenario in Figure 5.3, this procedure goes as follows:

- $Q_1 = \{$`Tom was broke`$\}$, so the text becomes 'Tom was broke.'

- $Q_2' = \{$`Tom broke a window/forced a door/picked a lock & went into the house`, `Tom stole a laptop`$\}$. To obtain Q_2 from Q_2', the node `Tom stole a laptop` is removed, so $Q_2 = \{$`Tom broke a window/forced a door/picked a lock & went into the house`$\}$. The text becomes 'Tom was broke. Then Tom broke a window/forced a door/picked a lock & went into the house.'

- $Q_3 = Q_3' = \{$`Tom stole a laptop`$\}$, so the text becomes 'Tom was broke. Then Tom broke a window/forced a door/picked a lock & went into the house. Then Tom stole a laptop.'

We furthermore propose to use the identifying scenario node to indicate which scenario was translated, in the form of 'Scenario node: text of scenario'.

Example (Adding the identifying scenario node to the text). The scenario text from the previous example is reported with the addition of the scenario node as follows:

> **Scenario: Tom was the burglar**: 'Tom was broke. Then, Tom broke a window/forced a door/picked a lock & went into the house. Then, Tom stole a laptop.'

Subscenarios and variations In combining the text of a scenario with (possibly multiple) subscenarios or variations we would like to capture the unfolding properties of a scenario with subscenarios and variations. We propose to consider subscenarios and variations as more elaborate versions of elements in a scenario. The main scenario is reported as separate text and additional texts specify elements of that main scenario in further detail (based on the variations and subscenarios).

To compose the text for some variations, the propositions corresponding to variation nodes are simply conjoined with 'or'. Though the order is not an issue, we know there is always an ordering on them (by definition of the variation idiom) so they can be put into some order in the text. The following example shows how this is done for the variation in Figure 5.1:

Example (Translating variations to text). The variation part of Figure 5.1 is shown in Figure 5.4. To translate these variations to text, they are conjoined with 'or', and the disjunction node is added as an identifier. This yields the following result: Tom broke a window/forced a door/picked a lock & went into the house: 'Tom forced a door & went into the house or Tom broke a window & went into the house or Tom picked a lock & went into the house'.

To compose the text for a subscenario, the same procedure as for scenarios can be used, as shown in the following example:

Example (Translating a subscenario to text). The subscenario portion of the example in Figure 5.1 is shown in Figure 5.5). With the procedure from the previous paragraph, the following text can be obtained:

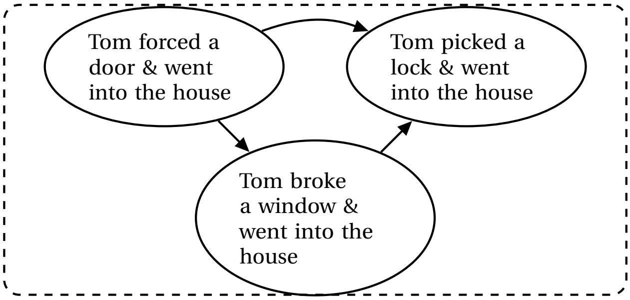


Figure 5.4: The variation structure from Figure 5.1.

> **Tom broke a window & went into the house**: A loose brick was near a house. Then, Tom threw brick at window. Therefore, the window broke. Then, Sylvia and Tom went into the house.

These texts for variations and subscenarios are now reported in addition to the main scenario. Note that the identifying text from a subscenario node or a disjunction nodes is always also an element in a (sub)scenario. It is thus clearly visible for which element subscenarios or variations provide an elaborate version.

Example (A variation as elaborating text). The main scenario from the previous paragraph includes the element 'Tom broke a window/forced a door/picked a lock & went into the house'. As a separate text, the detailed version of this element (with several variations) can be reported as

> **Tom broke a window/forced a door/picked a lock & went into the house**: 'Tom forced a door & went into the house or Tom broke a window & went into the house or Tom picked a lock & went into the house'.

This text relates to the element of the scenario which it elaborates on since it includes the text of the disjunction node as an identifier.

The full network structure shown in Figure 5.1 is thus translated as follows:

> **Scenario: Tom was the burglar**: 'Tom was broke. Then, Tom broke a window/forced a door/picked a lock & went into the house. Then, Tom stole a laptop.'
>
> **Tom broke a window/forced a door/picked a lock & went into the house**: 'Tom forced a door & went into the house or Tom broke a window & went into the house or Tom picked a lock & went into the house'

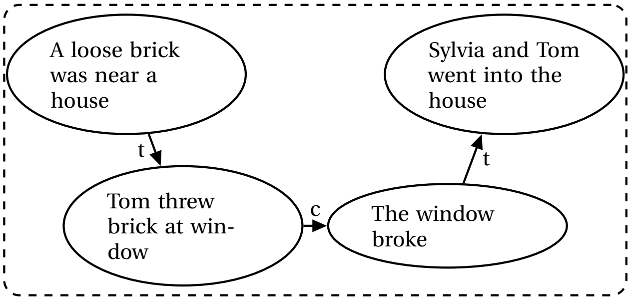


Figure 5.5: The subscenario structure from Figure 5.1.

Tom broke a window & went into the house: 'A loose brick was near a house. Then, Tom threw brick at window. Therefore, the window broke. Then, Sylvia and Tom went into the house.'

5.2 Reporting evidential support

Although the term 'evidential support' is used in a scenario-based approach, there is a clear interpretation of it in a probabilistic approach: evidential support is the change in probabilities that is the result of observing certain evidence (Carnap, 1962). However, how exactly this change in probabilities should be quantified in a numeric measure is a subject of debate. Though typical probabilistic approaches often report the strength of evidence in terms of likelihood ratios, our scenario-based approach calls for a different measure of evidential support, as will be explained below. In this section, we propose a numeric measure of evidential support and a verbal scale to interpret these numeric results qualitatively. This is used in the report in Section 5.4 to report for each piece of evidence and each scenario whether that evidence supports or attacks that scenario and how strong that support is.

A note about the likelihood ratio In a typical probabilistic approach to analysing legal evidence, the likelihood ratio (LR) is used to report the strength of evidence (Taroni et al., 2006). The likelihood ratio is used to compare two hypotheses, h_1 and h_2 (where h_2 is often the negation of h_1), by taking the fraction of the likelihoods of the evidence e turning up as a result of each of these hypotheses:

$$\mathrm{LR} = \frac{\Pr(e|h_1)}{\Pr(e|h_2)}.$$

This fraction also occurs in Bayes' rule in odds form (see Section 2.1) and when multiplying the prior odds of the two hypotheses with the likelihood ratio, the posterior odds are obtained.

The likelihood ratio is most informative when analysing two mutually exclusive and jointly exhaustive hypotheses (such as a hypothesis h and its negation) (Fenton et al., 2014). It is not an equally appropriate measure of evidential support for our current purposes. Firstly, in a scenario-based approach, there may be more than two alternative scenarios and they need not be jointly exhaustive. Secondly, alternative scenarios should be compared on scenario quality as well as evidential support, which is why we aim to report evidential support for each scenario separately such that it can later be taken into account when comparing scenarios.

Quantifying evidential support Our goal is to quantify the evidential support of a set of evidence e for a scenario with scenario node `ScN`. To measure evidential support, what matters is how $\Pr(\texttt{ScN = T})$ changes as a result of observing evidence e. We are thus interested in $\Pr(\texttt{ScN = T}|e)$ relative to $\Pr(\texttt{ScN = T})$. We propose to use the following measure for evidential support:

$$\frac{\Pr(\texttt{ScN = T}|e)}{\Pr(\texttt{ScN = T})}.$$

Other measures of evidential support exist which also compare how the posterior probability changes with respect to the prior probability. For example, $\frac{\Pr(e|\texttt{ScN = T})}{\Pr(e)}$ is equal to the measure proposed above, as can be found using Bayes' rule.

Definition of evidential support Using the measure proposed above, we have the following definition of evidential support, and the notions of supporting, attacking and neutral evidence:

Definition 5.2.1 (Evidential support)**.** For a scenario with scenario node `ScN` and evidence e, the *evidential support* of e for `ScN` is given by

$$\frac{\Pr(\texttt{ScN = T}|e)}{\Pr(\texttt{ScN = T})}.$$

Evidence e is

- supporting evidence if
$$\frac{\Pr(\texttt{ScN = T}|e)}{\Pr(\texttt{ScN = T})} > 1$$
- attacking evidence if
$$\frac{\Pr(\texttt{ScN = T}|e)}{\Pr(\texttt{ScN = T})} < 1$$

Table 5.1: A qualitative scale for evidential support.

	$x <$	0.001	Very strong evidence to attack
0.001	$\leq x <$	0.01	Strong evidence to attack
0.01	$\leq x <$	0.1	Moderate evidence to attack
0.1	$\leq x <$	1	Weak evidence to attack
1	$< x \leq$	10	Weak evidence to support
10	$< x \leq$	100	Moderate evidence to support
100	$< x \leq$	1000	Strong evidence to support
1000	$< x$		Very strong evidence to support

- neutral evidence if

$$\frac{\Pr(\texttt{ScN = T}|e)}{\Pr(\texttt{ScN = T})} = 1.$$

Evidential support for separate and combined evidence Evidence e in the definition of evidential support can be a single piece of evidence, or a set of evidence. To provide some insight into the case, we propose to report the evidential support of each piece of evidence for each scenario, but also the combined support of the collection of evidence for each scenario.

A verbal report of evidential support To qualify the amount of evidential support or attack by a piece or a set of evidence, we propose to use a verbal scale to translate the evidential support as calculated with the measure above. Similar scales are often used for the reporting of likelihood ratios, since likelihood ratios (and arguably our measure of evidential support as well) can be hard to interpret for a judge or jury (Nordgaard et al., 2012). Our proposed verbal scale is shown in Table 5.1. This scale was adapted from the standard quantitative scale published by the Association of Forensic Science Providers (2009). The original scale was intended for likelihood ratios associated with forensic evidence such as DNA matching, which is often associated with very high likelihood ratios. To accommodate for our use of the scale with evidential support, we adapted the scale to work with slightly less extreme numbers and to include translations below 1 (which were not included on the original scale).

5.3 Reporting scenario quality

A scenario-based approach to reasoning with legal evidence requires that scenarios are compared on scenario quality as well as evidential support. Explaining a Bayesian network in terms of scenarios thus requires that some information about

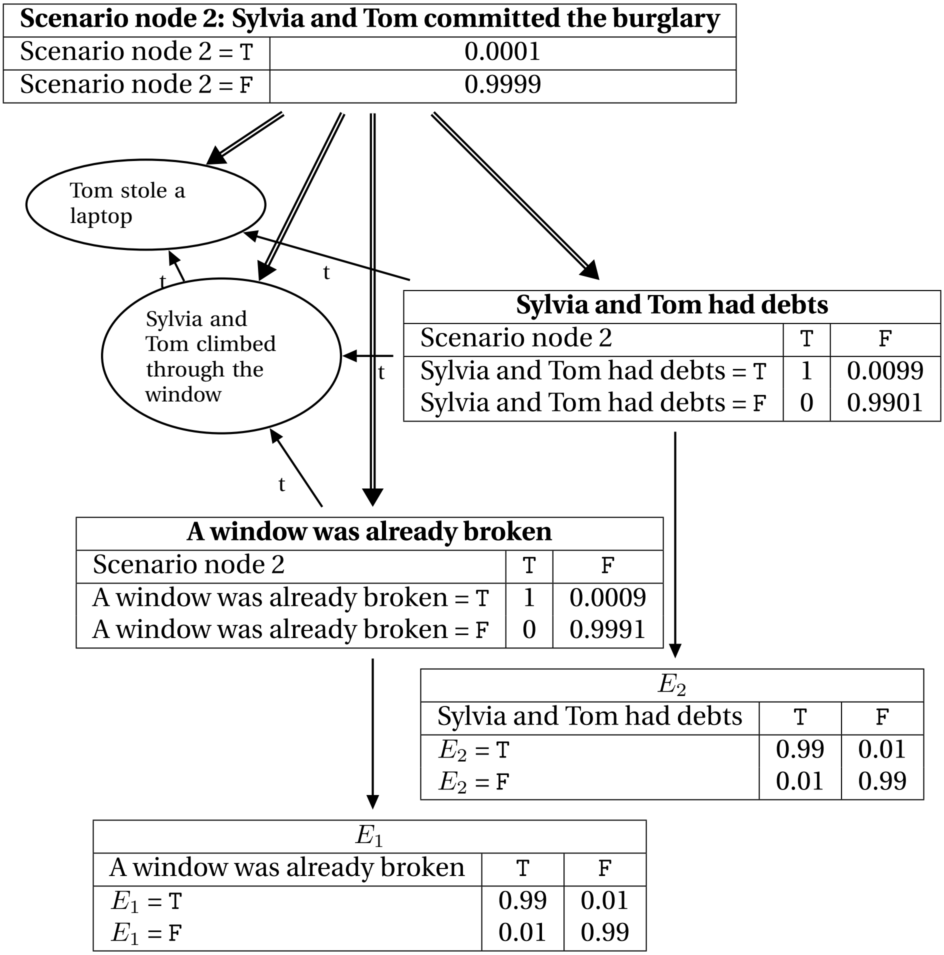


Figure 5.6: The second scenario from the burglary example of Chapter 3.

the quality of each scenario is communicated. In Chapter 3, the quality of a scenario was represented in a Bayesian network in terms of completeness, consistency and plausibility. In this section, we discuss how each of these three factors can be explained based on a Bayesian network as built with the method from Chapter 3. As will be shown in this section, the concept of plausibility is of particular interest for understanding the results of a Bayesian network, since the plausibility of an element of a scenario determines the impact of evidence for that element. To be precise, when an element of low plausibility is supported with evidence, the evidential support is stronger than that of a similar piece of evidence supporting a more plausible element. This will be illustrated with an example below.

Completeness A scenario was defined to be complete when there is some scenario scheme idiom which the scenario fits and completes. As was also discussed in Chapter 3, during the construction of a network, only complete scenarios should be represented. As a result, when explaining a Bayesian network constructed with the method from Chapter 3, all scenarios are expected to be complete and reported as such. Nonetheless, their completeness can be checked against the collection of scenario scheme idioms (or those developed on the fly by the modeller) and when they are incomplete, this is reported.

Consistency The consistency of a scenario depends on whether it contained any elements that are jointly inconsistent. Elements are jointly inconsistent when the probability of them occurring together is 0. To model such an inconsistency, a modeller can use a constraint node construction, as discussed in Chapter 3. To report about the consistency of a scenario, it can thus be checked whether there is such a constraint node on two or more elements of the scenario, leading to a probability of 0 for the events occurring together. In this case, the scenario is reported as being inconsistent.

Plausibility The plausibility of a scenario can be read from the prior probability of the scenario node. However, to explain the plausibility of a scenario, it is best to explain the plausibility of elements in that scenario. The plausibility of an element was defined as the probability of that element occurring, possibly conditioned on a parent node. To explain the plausibility of elements, we aim to report any implausible elements a scenario may include.

A threshold for implausibility While we defined completeness and consistency as boolean notions, with a scenario being either complete or not and either consistent or not, plausibility can be a gradual notion. Reporting on implausible elements therefore requires some threshold to determine when an element goes from being plausible to being implausible. We propose to use a threshold of 0.01, meaning that an element with a plausibility below 0.01 is reported as being implausible. Depending on the application, a different threshold can be chosen.

The impact of evidence for implausible elements When an element in a scenario is considered implausible, it can still become probable when enough evidence is given to support it. This corresponds to the point made by Wagenaar et al. (1993), that even a very implausible scenario can be believable when there is enough evidence for it. Moreover, we observe that evidence for an implausible element has a greater impact on the probability of the whole scenario. Consider as an example the second scenario that was modelled in Chapter 3, shown in Figure 5.6 with some probabilities specified and some evidence E_1 connected to `A window was already broken` and E_2 connected to `Sylvia and Tom had debts`. The numbers in the probability table of element `A window was already broken` were calculated to represent a plausibility of 0.001 and the probabilities for `Sylvia and Tom had debts` represent a plausibility of 0.01.

With these numbers, evidence e_1 (short hand for E_1 = `T`) lends greater evidential support to the scenario node than evidence e_2 (short hand for E_2 = `T`). The evidential support of e_1 is 90, while the evidential support of e_2 is 50. This difference results from the difference in plausibility between the events supported by e_1 and e_2 (since the probability tables of e_1 and e_2 are equal). To see why this is the case, consider the term $\Pr(e_1)$ in the denominator when calculating the posterior probability of the scenario node given evidence e_1 using Bayes' rule (a similar calculation is done for e_2):

$$\Pr(\texttt{Scenario node 2 = T}|e_1) = \frac{\Pr(e_1|\texttt{Scenario node 2 = T}) \cdot \Pr(\texttt{Scenario node 2 = T})}{\Pr(e_1)}.$$

Due to the implausibility of the element `A window was already broken`, the probability of the evidence being observed $\Pr(e_1)$ is low, as follows from the following calculation (assuming that the probability of the evidence occurring if the window is not broken is low):

$$\begin{aligned}\Pr(e_1) = &\Pr(e_1|\texttt{A window was already broken = T})\\ &\cdot \Pr(\texttt{A window was already broken = T})\\ &+ \Pr(e_1|\texttt{A window was already broken = F})\\ &\cdot \Pr(\texttt{A window was already broken = F})\\ &= 0.01.\end{aligned}$$

Since the plausibility of `Sylvia and Tom had debts` is somewhat higher, a similar calculation leads to the result that $\Pr(e_2) = 0.02$. Using these numbers in Bayes' rule above, means that the probability of the scenario node increases more on evidence e_1 than on e_2, due to a lower number in the denominator. In the absence of other differences, evidence for an element with lower plausibility thus leads to higher evidential support than evidence for an element with higher plausibility.

Evidential gaps As shown above, implausible elements can result in a great effect on the whole scenario when they are supported by evidence. Simultaneously, when implausible elements are unsupported, they are the 'weakest links' in a scenario and are thus worth pointing out, especially because if evidence were found for these weakest links, the entire scenario could possibly become much more believable. Such crucial parts of a scenario that require evidence for a scenario to become believable are called evidential gaps (Bex, 2011). To explain the plausibility of a scenario, we propose to report implausible elements when they are unsupported as evidential gaps and when they are supported as supported implausible elements.

5.4 A reporting format

With the material from the previous sections, information about the scenarios in a Bayesian network can be reported. The result is an explanation that includes all ingredients needed for a scenario-based assessment of the case: (1) the alternative scenarios, (2) their quality and (3) the evidential support for each scenario. Accompanying these ingredients with probabilistic information, we propose that a report about a Bayesian network should consist of the following, for each scenario:

1. the scenario in text form with its prior and posterior probability; and
2. whether the scenario is complete and consistent and a list of implausible elements in the scenario, including whether they are supported by evidence or remain evidential gaps; and
3. evidential support of each piece of evidence for that scenario and a combined measure of evidential support (of the collection of all evidence).

Such a report can potentially be produced automatically. A scenario in text form can be produced with the method from Section 5.1, and the prior and posterior probability of the scenario can be read from the network (at the scenario node corresponding to that scenario).

Whether a scenario is complete can be determined by comparing it to a scenario scheme idiom to find whether the scenario fits and completes that scheme idioms. Whether a scenario is consistent can be found by checking if there is a constraint node on elements of that scenario which results in the probability of these elements together occurring to be 0.

Implausible elements can be found in the network as elements with a probability below 0.01 as explained in Section 5.3. When an implausible element has no evidence directly connected to it, it is reported as an evidential gap. When there is evidence for it, it is pointed out as an implausible element with support.

For each piece of evidence, the evidential support of that evidence for the scenario can be calculated with the measure from Section 5.2 and translated to a verbal account with the verbal scale from that same section. Similarly, a combined measure of evidential support is calculated and reported verbally.

An example As an illustration of the aforementioned reporting format, reconsider the burglary example from Chapter 3. As an example, some evidential nodes were added as shown in Figure 5.7. These evidential nodes are the observation of a broken window (`Broken window`), a statement of someone stating that Tom tried to sell them a laptop (`Statement: Tom sold laptop`) and a testimony of someone who saw that the window was already broken before the burglary (`Testimony: window was already broken`). Probabilities were specified only for illustrative purposes, available as a GenIe model on www.charlottevlek.nl/networks. The result is the following report:

- **Scenarios in the network**:
  - Scenario 1: Tom was the burglar (prior probability: 0.0001, posterior probability: 0.0133):

    Scenario: Tom was the burglar: Tom was broke. Then, Tom broke a window/forced a door/picked a lock & went into the house. Then, Tom stole a laptop.

    Tom broke a window/forced a door/picked a lock & went into the house: Tom forced a door & went into the house or Tom broke a window & went into the house or Tom picked a lock & went into the house.

    Tom broke a window & went into the house: A loose brick was near a house. Then, Tom threw brick at window. Therefore, the window broke. Then, Tom went into the house.
  - Scenario 2: Sylvia and Tom committed the burglary. (prior probability: 0.0001, posterior probability: 0.2326)

    Scenario: Sylvia and Tom committed the burglary: Sylvia and Tom had debts and a window was already broken. Then, Sylvia and Tom climbed through the window. Then, Tom stole a laptop.
- **Scenario quality**

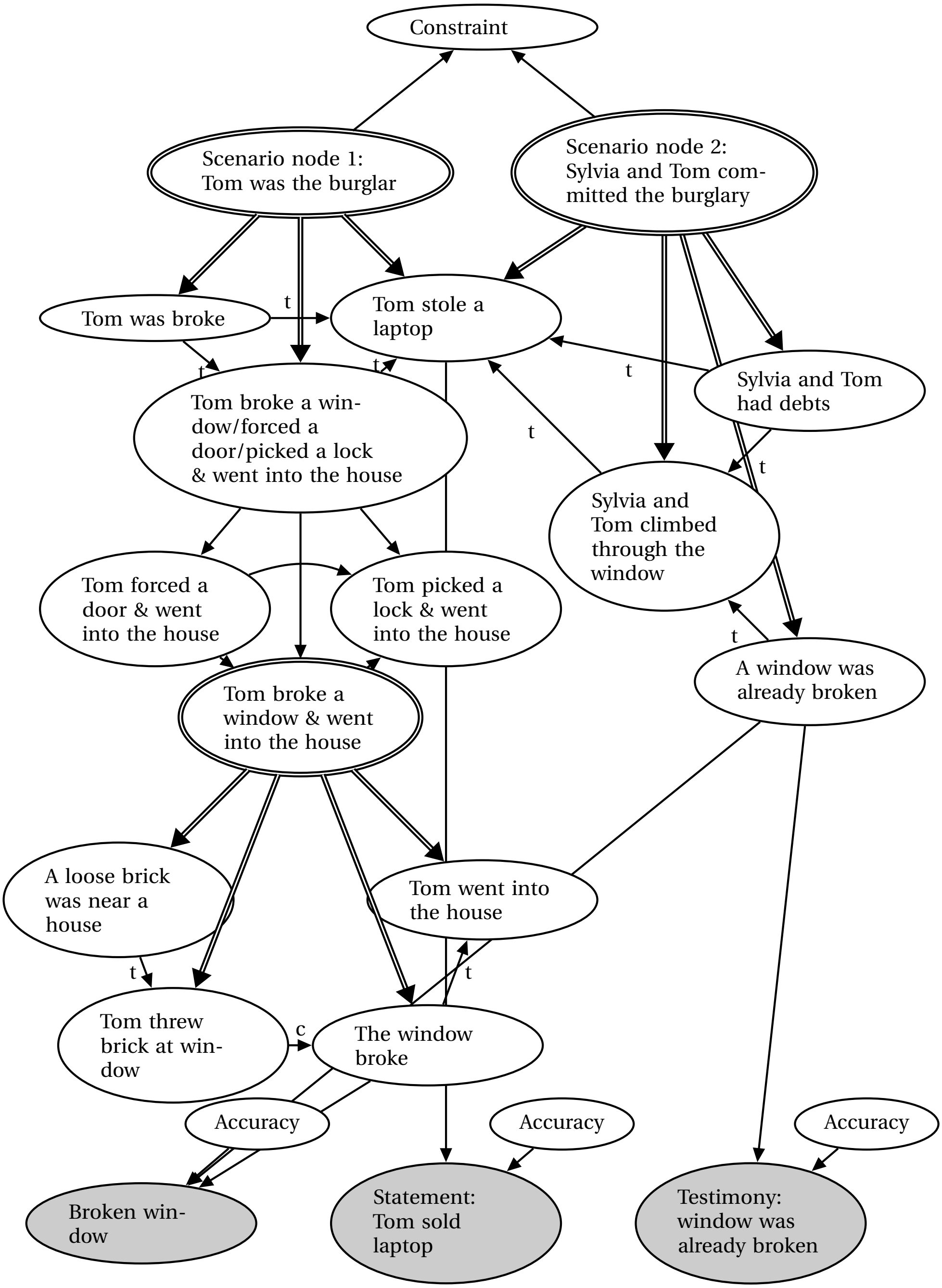


Figure 5.7: The burglary example from Chapter 3, with some evidential nodes (indicated as grey nodes) added for illustrative purposes.

  - Scenario 1 is complete and consistent. It contains the evidential gap ‘Tom was broke’.
  - Scenario 2 is complete and consistent. It contains the evidential gap ‘Sylvia and Tom had debts’ and the supported implausible element ‘A window was already broken’.

- **Evidence related to each scenario**
  - Evidence for and against scenario 1:
    * Broken window: moderate evidence to support scenario 1.
    * Statement: Tom sold laptop: moderate evidence to support scenario 1.
    * Testimony: window was already broken: weak evidence to attack scenario 1.
    * All evidence combined: strong evidence to support scenario 1.
  - Evidence for and against scenario 2:
    * Broken window: moderate evidence to support scenario 2.
    * Statement: Tom sold laptop: moderate evidence to support scenario 2.
    * Testimony: window was already broken: weak evidence to support scenario 2.
    * All evidence combined: very strong evidence to support scenario 2.

5.5 Conclusion

In this chapter, we proposed to explain a Bayesian network in terms of scenarios. Since the techniques proposed in this chapter employ properties from the construction method from Chapter 3, these explanation techniques are considered an extension of that method, resulting in one method for constructing and explaining Bayesian networks with scenarios.

Our aim in this chapter was to explain a Bayesian network with scenarios in such a way that the a judge or juror is presented with all the ingredients needed for a scenario-based analysis of a case, namely: (1) the alternative scenarios, (2) their quality and (3) the evidential support for each scenario. Simultaneously, this information gives feedback on the modelling choices that went into the model, particularly the content of the network (the alternative scenarios), how the quality of each scenario was represented and how the evidence was related to them.

To enable a report about these three components needed for a scenario-based approach, we first proposed a technique for extracting the alternative scenarios that are represented in a Bayesian network and translate them to text. In the report, these texts are accompanied with probabilistic information, namely the prior and posterior probability of each scenario.

Secondly, a measure of evidential support was proposed. For each scenario and each piece of evidence, the evidential support is calculated and reported qualitatively using a verbal scale. The evidential support of the collection of all evidence in the case is also reported qualitatively.

Thirdly, the quality of a scenario as represented in a network was treated. In the resulting report, it is explained whether a scenario is complete and consistent. Special attention was directed towards plausibility, since the implausibility of elements in a scenario can help to understand why a piece of evidence connected to an implausible part of a scenario results in higher evidential support. On the one hand, implausible elements with evidence supporting them are of interest since they lead to higher evidential support, while on the other hand implausible elements that remain unsupported point out crucial weaknesses in a scenario, known as evidential gaps. In the report, evidential gaps as well as implausible elements with support are reported.

This method amounts to an extension of the construction method to explain to a judge or jury what has been modelled in the Bayesian network. Moreover, with the information from the report, a judge or jury can now analyse the case with a scenario-based approach, but enhanced with qualitative probabilistic information from the network.

6. Case study: the Nijmegen case

In this chapter, we perform a case study to test our extended method with explanation techniques proposed in Chapter 5. We first construct a network for the case with the method from Chapter 3, since this is required for the application of the explanation techniques. However, the main focus of this chapter is on evaluating the explanation techniques from Chapter 5. We use the following criteria to evaluate the method:

1. Does the report give insight into the content of the model?
2. Does the report give insight into what led to the evidence?
3. Does the report give insight into the conclusions of the Bayesian network?
4. Does the extraction of scenarios produce informative results for the case at hand?
5. Does the reporting of scenario quality and evidential support produce informative results for the case at hand?

The first three criteria were inspired by the discussion of explanation methods by Lacave and Díez (2002), who state that the content of an explanation should include three things: the content of the model, an explanation of the evidence and the reasoning in the model. Our method aims to explain the content of the model in terms of scenarios, therefore extracting scenarios and reporting on their quality to explain what was modelled in the network. With the first criterion, we aim to evaluate how well this communicates the content of the model. The reported scenarios also serve as explanations of the evidence, and in the second criterion it is evaluated how well these scenarios give insight into how the evidence came to be. Finally, with the reporting of evidential support and possibly evidential gaps, we aim to provide insight into the conclusions of the Bayesian network, which is evaluated in the third criterion.

The last two criteria concern the specific features of our proposed explanation techniques. We aim to evaluate our scenario extraction procedure by finding whether it produces good results for a real case (criterion 4). We furthermore aim to evaluate whether the reporting of scenario quality and evidential support produces the expected results when applied to a real case (criterion 5).

The remainder of this chapter is organised as follows: first we briefly discuss the case and why it is of interest for a case study (Section 6.1). Then we construct a Bayesian network for this case (Section 6.2). Then a report is compiled for the case (Section 6.3). Finally, the results are evaluated according to the aforementioned criteria (Section 6.4).

6.1 The case

This case is about a suspect who is thought to have helped to move a dead body. The case took place in the city of Nijmegen in The Netherlands, which is why we refer to it as 'the Nijmegen case'. The suspect is not accused of killing the victim, but of helping the killer move the body to a location out of town, after the killing took place in a house in town. The suspect was convicted for his help in moving the body, but later the Court of Appeals concluded that there was not enough evidence for a conviction, so the suspect was found innocent. The appeal case can be found (in Dutch) on www.rechtspraak.nl using code ECLI:NL:GHARL:2014:8941.

This case is of interest for testing our method because it includes some probabilistic reasoning as well as several alternative scenarios about what may have happened. The probabilistic reasoning (though non-numerical) is concerned with DNA-evidence in this case. DNA traces matching with the suspect were found on the victim's body. Several alternative activity-level hypotheses were compared: the suspect moved the body, thereby leaving traces, or the suspect was not involved in moving the body but his DNA was transferred from a couch in the victim's home. In the appeal case, another hypothesis was added, namely that the suspect was not involved in moving the body, but his DNA was transferred from a blanket (used by the suspect some time prior to the killing), which was used by the killer to wrap the body and transport it to another location. For each of these hypotheses, the likelihood of finding DNA traces given each hypothesis was compared qualitatively in the appeal case. In our analysis of the case, we will provide a numerical interpretation of this qualitative discussion in our model.

In what follows, we use fake names for the suspect (Adam), the supposed killer (Bert) and the victim (Chris).

6.2 Constructing the network

6.2.1 Scenario 1

The first scenario in this case is as follows: Adam, Bert and Chris all knew each other because they were involved in a cannabis operation. Bert killed Chris and then Adam helped Bert to carry Chris's body to a car. They drove to the countryside and dumped the body there.

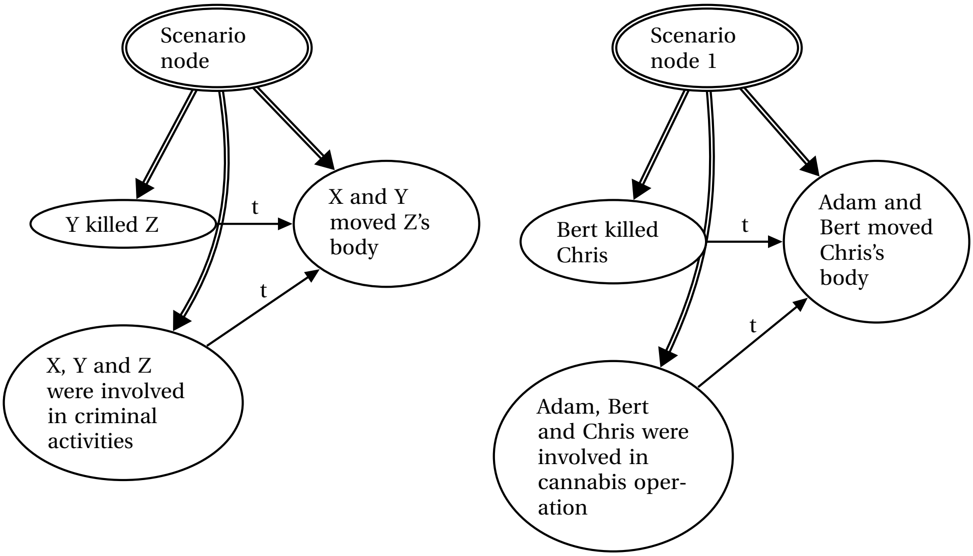


Figure 6.1: A scenario scheme idiom (left) and the initial scenario structure for scenario 1 (right).

A scenario scheme idiom for this scenario is shown in Figure 6.1 (left), as well as the scenario as modelled using this scheme (right). In this specific case, details about how they moved the body are of interest. The element 'Adam and Bert moved Chris's body' is thus unfolded to a subscenario, as shown in Figure 6.2. Furthermore, the element 'Adam, Bert and Chris were involved in a cannabis operation' is also unfolded, since it will become important in relation to the other scenarios (see below).

6.2.2 Scenario 2

In the second scenario, Bert worked alone and also moved the body by himself. Note that this is effectively an innocent scenario for the current suspect Adam. This scenario can be represented with a scenario scheme with the same structure as the one that was used earlier, but with a node 'Y moved Z's body' instead of 'X and Y moved Z's body'. This second scenario also includes the element 'Adam, Bert and Chris were involved in a cannabis operation', because this element (when worked out as a subscenario as shown in Figure 6.2) helps to explain why evidence was found that related the crime to Adam: DNA traces of Adam were on the couch at the cannabis plantation and then transferred to Chris's body. A network with both scenarios is shown in Figure 6.3.

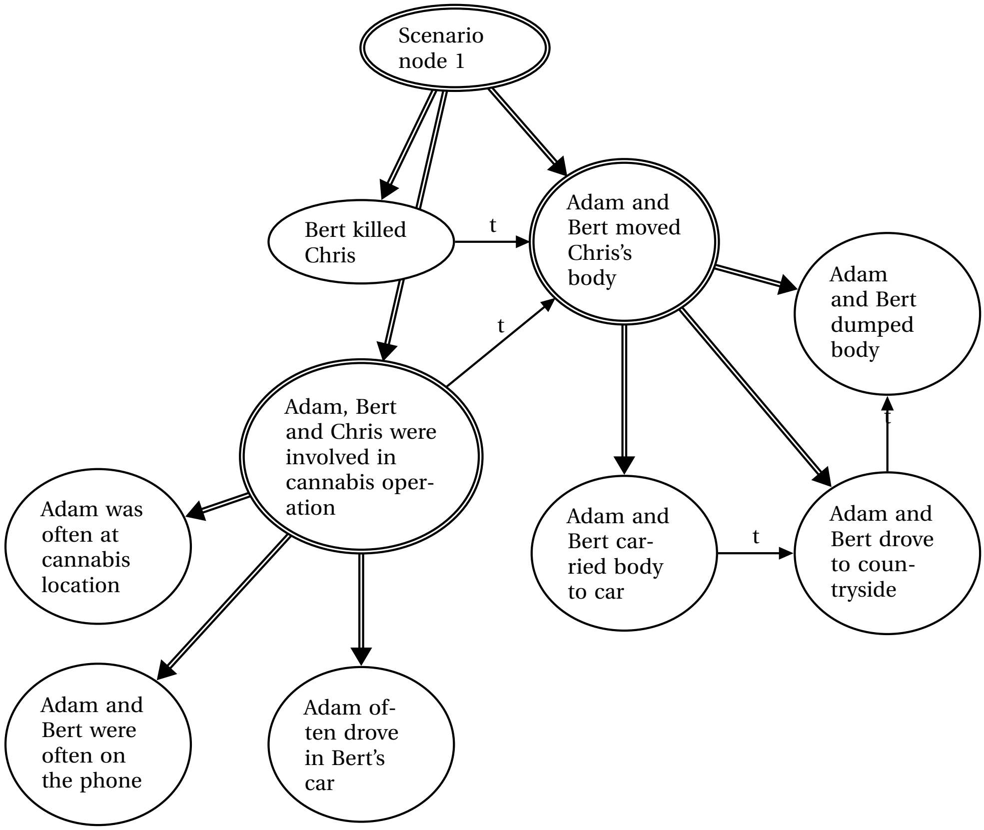


Figure 6.2: The first scenario with subscenarios.

6.2.3 Scenario 3

In the appeal case, a third scenario was considered in addition to the two scenarios described above. In this scenario, Bert also worked alone and moved the body by himself, but now he used a blanket to move the body and this blanket contained traces of Adam, because Adam sometimes stayed at the cannabis plantation during the night. As we will see later on, the probability of a transfer of Adam's DNA is now higher than when transfer took place via the couch. This third scenario is represented with the same scenario scheme as the previous (second) scenario. A network with all three scenarios is shown in Figure 6.4.

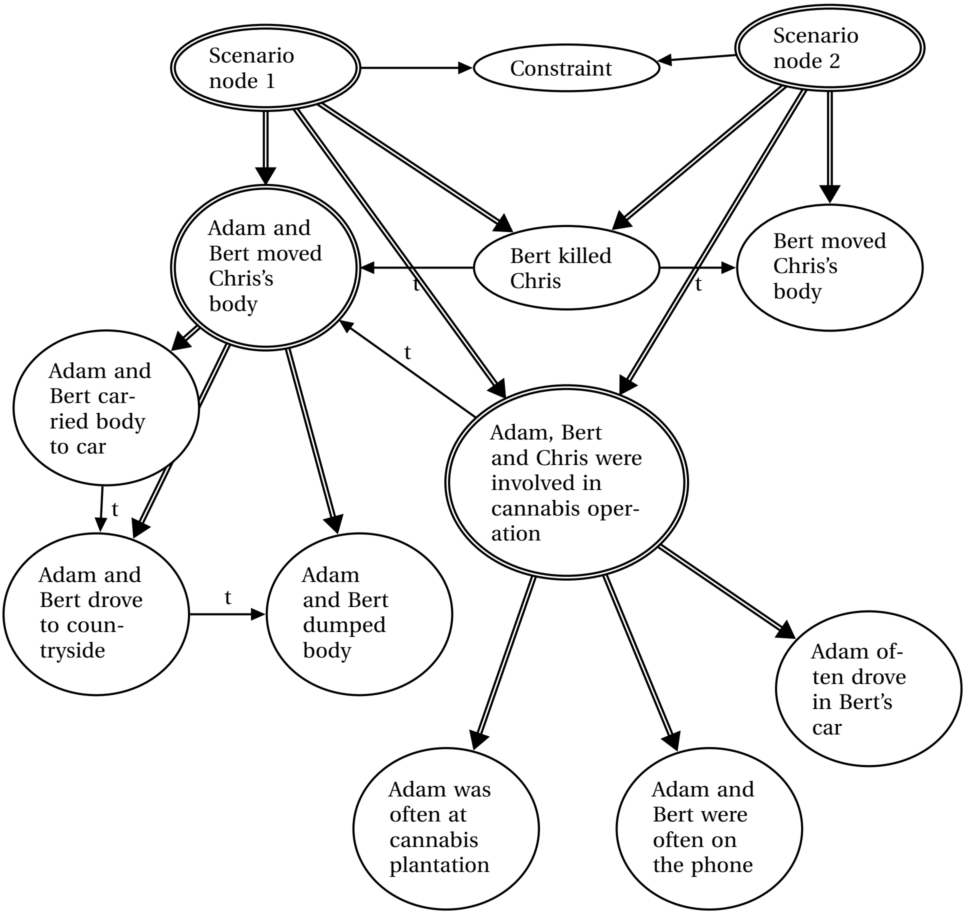


Figure 6.3: The first two scenarios in one network.

6.2.4 The evidence

Evidence in this case is the following:

- Adam's car was not seen on any surveillance camera's (`Adam's car on ARS cameras`)
- A DNA match was found between Adam's profile and traces on Chris's body (`DNA match`)
- A hair matching Adam's was found on some duct tape (`Hair on duct tape`)

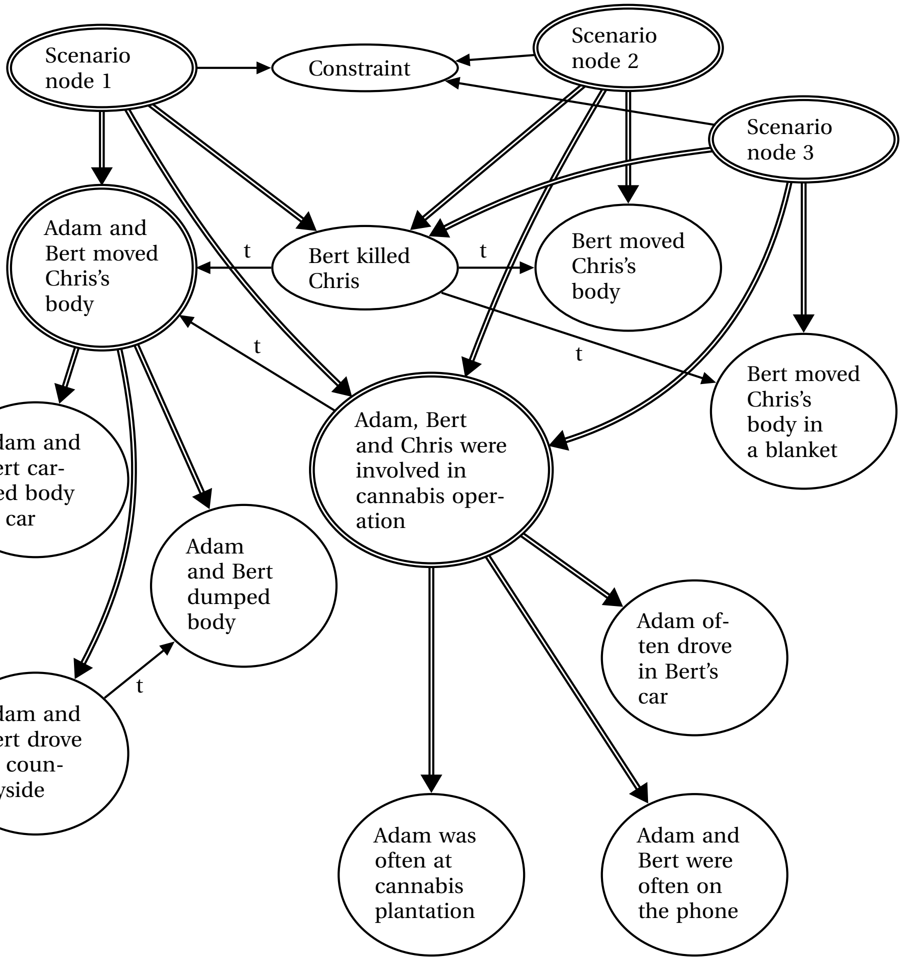


Figure 6.4: The three scenarios in one network.

- Bert was already convicted for the killing of Chris (`Bert's conviction`)
- Chris's body was found in the countryside (`Body in countryside`)
- Adam and Bert talked on the phone several times at the night of the killing (`Phone calls Adam and Bert`)
- Traces in Bert's car left by Adam (`Traces of Adam in car`)

Table 6.1: The probability table for `DNA match`.

`Adam and Bert carried body`	T				F			
`Bert moved body in blanket`	T		F		T		F	
`Adam often at cann. plant.`	T	F	T	F	T	F	T	F
`DNA match = T`	0.99	0.99	0.99	0.99	0.9	0.9	0.1	0
`DNA match = F`	0.01	0.01	0.01	0.01	0.1	0.1	0.9	1

Each piece of evidence can be represented as a separate node in the network. This leads to a graph as shown in Figure 6.5.

6.2.5 The probabilities

A probabilistic discussion of the DNA match played an important role in the appeal case. We capture the qualitative analysis in the numbers of the probability table for the node `DNA match`. In the initial case, the court considered only two activity-level hypotheses: Adam and Bert carried the body together, or Bert moved Chris's body by himself and Adam's DNA was transferred some other way (because Adam was often at the cannabis plantation). An expert stated that the DNA evidence would be 'much more likely' to be found under the hypothesis that Adam and Bert carried the body than under the hypothesis that Bert moved Chris's body by himself and the DNA was transferred some other way. In the appeal case, a third hypothesis was considered, namely that Bert moved the body in a blanket which Adam had used previously when he was sleeping at the cannabis plantation. From the court records it becomes apparent that given this third hypothesis, the DNA evidence is still considered quite likely, although it is unclear how it compares to the other two hypotheses. In the probability table of `DNA match`, these considerations can be quantified with some interpretation. We set the probabilities as shown in Table 6.1, such that

- if `Adam and Bert carried body to car = T` the probability of finding the DNA match is 0.99;
- if `Adam and Bert carried body to car = F` and `Bert moved Chris's body in a blanket = T`, the probability of a DNA match is 0.9;
- if `Adam and Bert carried body to car = F` and `Bert moved Chris's body in a blanket = F` with `Adam was often at cannabis plantation = T` the probability of finding a DNA match is 0.1;
- if `Adam and Bert carried body to car = F` and `Bert moved Chris's body in a blanket = F` and `Adam was often at cannabis plantation = F`, the probability of a DNA match is 0.

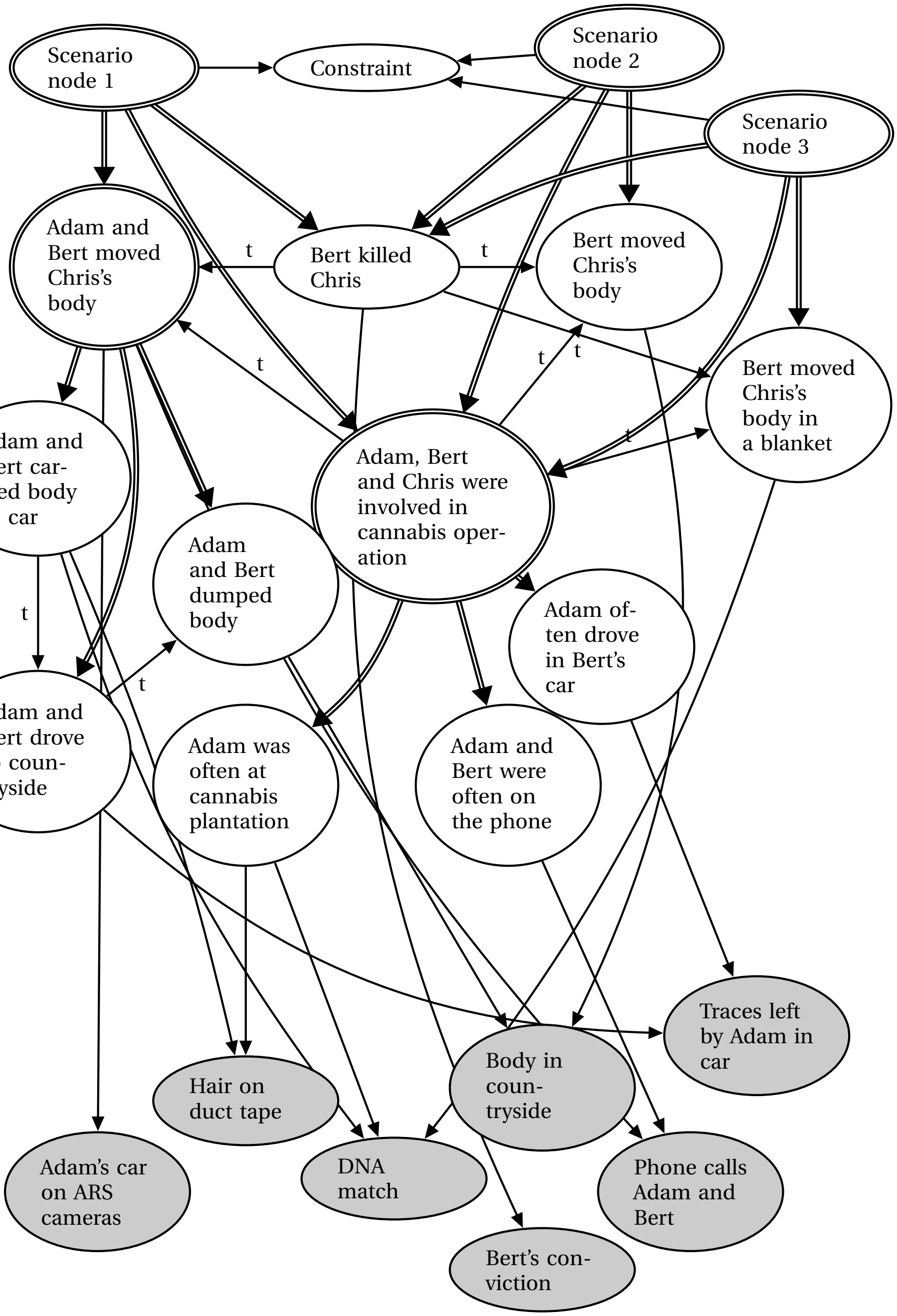


Figure 6.5: A network for the case study: The three scenarios with evidence. Evidential nodes are indicated as grey nodes.

This way, the DNA evidence is much more likely under the hypothesis that Adam and Bert carried the body (probability of 0.99) than under the hypothesis that it was transferred simply because Adam was often at the cannabis plantation (probability of 0.1). In contrast, the probability of finding DNA evidence under the hypothesis that Bert moved the body in a blanket (probability of 0.9) is closer to the probability that resulted from the hypothesis that Adam and Bert carried the body. This reflects the decision in the appeal case that the initial case should have taken the third explanation of the DNA traces into account.

For the assessment of other numbers, elicitation techniques exist, see, for instance, Renooij (2001). The scenario nodes are a special case, since there the prior probability is expressed that each scenario as a whole is true. For lack of further information about this, we propose to set the probability of each scenario to be true to the low probability of 0.001, well below our threshold of 0.01 for plausibility. Furthermore, numbers in the probability tables of the network in Figure 6.5 are partially constrained due to the scenario scheme idioms. The probability of an element of a scenario conditioned on the scenario node being true is always one: $\Pr(\text{Element} = \texttt{True}|\texttt{Scenario node = T}) = 1$ and similarly for elements of a subscenario (such as $\Pr(\texttt{Adam and Bert carried body to car = T}|\texttt{Adam and Bert moved Chris's body = T}) = 1$).

The assessment of probabilities for elements of the scenarios needs to take into account plausibility, as explained in Chapter 3. For example, the element that Bert killed Chris seems implausible and we would like to set the numbers such that $\Pr(\texttt{Bert killed Chris = T}) = 0.01$. To calculate the numbers as they should appear in the probability table, we use

$$
\begin{aligned}
0.01 &= \Pr(\texttt{Bert killed Chris = T}) \\
&= \Pr(\texttt{Bert killed Chris = T}|\texttt{ScN 1 = T}) \cdot \Pr(\texttt{ScN 1 = T}) \\
&+ \Pr(\texttt{Bert killed Chris = T}|\texttt{ScN 1 = F}) \cdot \Pr(\texttt{ScN 1 = F}) \\
&= 1 \cdot \Pr(\texttt{ScN 1 = T}) \\
&+ \Pr(\texttt{Bert killed Chris = T}|\texttt{ScN 1 = F}) \cdot \Pr(\texttt{ScN 1 = F}) \\
&= 1 \cdot 0.001 + \Pr(\texttt{Bert killed Chris = T}|\texttt{ScN 1 = F}) \cdot 0.999
\end{aligned}
$$

so the result of $\Pr(\texttt{Bert killed Chris = T}) = 0.01$ is obtained by setting in the probability table:

$$\Pr(\texttt{Bert killed Chris = T}|\texttt{ScN 1 = F}) = \frac{0.01 - 0.001}{0.999} \approx 0.009.$$

As we can see, due to the low prior for the scenario node, this number approximates the value of $\Pr(\texttt{Bert killed Chris = T})$. However, with other prior probabilities, they may differ more.

In a similar fashion, other numbers for the probability tables can be elicited by going through the nodes starting at the roots (without any parents), then their child nodes, et cetera. For example, after eliciting the tables of `ScN1`, `Bert killed Chris` and `Adam, Bert and Chris were involved in cannabis operation`, the conditional probabilities in the table of `Adam and Bert moved Chris's body` can be calculated. We estimate this element to be somewhat plausible (with a probability of 0.1), given that Bert killed Chris and they all knew each other:

$$\begin{aligned}&\Pr(\texttt{Adam and Bert moved Chris's body = T}|\texttt{Bert killed Chris = T},\\ &\quad\texttt{Adam, Bert and Chris were involved in cannabis operation = T})\\ &\quad= 0.1.\end{aligned}$$

A Bayesian network model is available at www.charlottevlek.nl/networks.

6.3 The report

With a Bayesian network available, a report can be compiled to explain the network. This report will consist of three parts, as discussed in Chapter 5:

1. The scenario in text form with its prior and posterior probability; and
2. Whether the scenario is complete and consistent and a list of implausible elements in the scenario, including whether they are supported by evidence or remain evidential gaps; and
3. Evidential support of each piece of evidence for that scenario and a combined measure of evidential support (of the collection of all evidence).

It is possible to extract this report fully automatically, since information needed for all three parts can be computed from the network directly. For illustration purposes, the following report was compiled manually for the network in Figure 6.5.

- **Scenarios in the network**:
  - Scenario 1 (prior probability: 0.001, posterior probability: 0.5296):

    Scenario: Bert killed Chris, and Adam, Bert and Chris were involved in cannabis operation. Then Adam and Bert moved Chris's body.

    Adam, Bert and Chris were involved in cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car.

Adam and Bert moved Chris's body: Adam and Bert carried body to car. Then Adam and Bert drove to countryside. Then Adam and Bert dumped body.

- Scenario 2 (prior probability: 0.001, posterior probability: 0.1180):

  Scenario: Bert killed Chris, and Adam, Bert and Chris were involved in cannabis operation. Then Bert moved Chris's body.

  Adam, Bert and Chris were involved in cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car.

- Scenario 3 (prior probability: 0.001, posterior probability: 0.2913):

  Scenario: Bert killed Chris, and Adam, Bert and Chris were involved in cannabis operation. Then Bert moved Chris's body in a blanket.

  Adam, Bert and Chris were involved in cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car.

- **Scenario quality**
  - Scenario 1 is complete and consistent. It contains the supported implausible element `Bert killed Chris`.
  - Scenario 2 is complete and consistent. It contains the supported implausible element `Bert killed Chris`.
  - Scenario 3 is complete and consistent. It contains the supported implausible element `Bert killed Chris`.

- **Evidence related to each scenario**
  - Evidence for and against scenario 1:
    * Adam's car not on ARS cameras: weak evidence to attack scenario 1.
    * DNA match: moderate evidence to support scenario 1.
    * Hair on duct tape: moderate evidence to support scenario 1.
    * Bert's conviction: moderate evidence to support scenario 1.
    * Body in countryside: strong evidence to support scenario 1.
    * Phone calls Adam and Bert: weak evidence to support scenario 1.
    * Traces of Adam in car: weak evidence to support scenario 1.
    * All evidence combined: strong evidence to support scenario 1.
  - Evidence for and against scenario 2:
    * Adam's car not on ARS cameras: weak evidence to attack scenario 2.
    * DNA match: moderate evidence to support scenario 2.

    * Hair on duct tape: moderate evidence to support scenario 2.
    * Bert's conviction: moderate evidence to support scenario 2.
    * Body in countryside: strong evidence to support scenario 2.
    * Phone calls Adam and Bert: weak evidence to support scenario 2.
    * Traces of Adam in car: weak evidence to support scenario 2.
    * All evidence combined: strong evidence to support scenario 2.
  - Evidence for and against scenario 3:
    * Adam's car not on ARS cameras: weak evidence to attack scenario 3.
    * DNA match: moderate evidence to support scenario 3.
    * Hair on duct tape: moderate evidence to support scenario 3.
    * Bert's conviction: moderate evidence to support scenario 3.
    * Body in countryside: strong evidence to support scenario 3.
    * Phone calls Adam and Bert: weak evidence to support scenario 3.
    * Traces of Adam in car: weak evidence to support scenario 3.
    * All evidence combined: strong evidence to support scenario 3.

6.4 Discussion of the case study

The case study in this chapter was meant to evaluate our proposed method for constructing and explaining Bayesian networks, with a focus on evaluating the explanation techniques. In this discussion, we present our findings concerning the criteria proposed in the introduction of this chapter.

6.4.1 Criterion 1: the content

By reporting which scenarios were modelled in the network and how their quality was represented, we aimed to report on the content of the model. For this case, the resulting report in Section 6.3 included the three alternative scenarios in which either Adam helped Bert move Chris's body, or Bert moved the body by himself, or Bert moved the body in a blanket. The report also included some remarks about the quality of each scenario and the evidence supporting each scenario, which gave a good impression of the scenarios, their quality and their evidential support as modelled in the network.

Although these three scenarios do communicate the content of the network, the report does not provide much insight into how the three scenarios were represented. Specifically, local modelling choices (such as which piece of evidence is connected to which (sub)hypothesis) are not communicated. This can be viewed as an advantage as well as a disadvantage. On the one hand, a judge or jury can now not really check whether the modelling choices are correct, in particular concerning the detailed choices of specific local connections between nodes. On the other hand, by not reporting on all the details but only on the high-level scenarios, a judge or jury can gain insight into the model without being overwhelmed by too much information.

6.4.2 Criterion 2: the evidence

In a scenario-based approach, the scenarios are viewed as explanations of the evidence. The alternative scenarios about Adam helping Bert move the body or Bert moving it by himself (possibly in a blanket) thus take the role of alternative explanations for the evidence as they can be found in the network.

These alternative scenarios give some insight into how the evidence came to be, though not explicitly. They reveal alternative explanations of the crucial pieces of evidence, namely why Adam's DNA was found on the victim and why there were several phone calls between Adam and Bert. However, in the remainder of the reporting format, scenarios are not communicated as being explanations of the evidence, but rather as being supported by evidence. Instead of a narrative perspective in which scenarios are explanations, our report adheres to a probabilistic perspective in which the hypotheses are supported by evidence. It could be an interesting topic for further study to find whether a different reporting format in which scenarios serve as explanations of the evidence would be more insightful for a judge or jury.

6.4.3 Criterion 3: the reasoning

By reporting evidential support of each piece of evidence for each scenario in combination with reporting any evidential gaps and implausible elements in the scenario, our method aimed to give insight into the change from prior to posterior probability of each scenario node. In the report for this case, we saw that there was a clear difference between the various pieces of evidence and their reported evidential support: the finding of a body in the countryside provided strong evidence to support scenario 1, while the phone calls provided only weak evidence. The reporting of the evidential support thus provides insight into how the effects of the evidence are incorporated in the resulting posterior probabilities.

When comparing the three scenarios and their list of evidence, we see that the reported evidential support of each piece of evidence is the same for all three scenarios. This is because the numerical evidential support of each piece of evidence, though not exactly the same, is in the same order of magnitude for all three scenarios. The report thus conveys the general sense that there is not much difference between the evidential support for the three scenarios, which matches the conclusion of the appeal case: though assigning different likelihoods to the DNA evidence as a result of the three competing hypotheses, it was concluded that the evidence did not sufficiently distinguish between the three alternative scenarios.

6.4.4 Criterion 4: extracting scenarios

To report on the alternative scenarios that are in the network, a technique for extracting scenarios was used. In this case study we found that the three scenarios could be adequately extracted from the network, producing a text that did indeed reflect each scenario as it was intended. A bit of redundancy can be found in the report, where one subscenario was reported three times, since it was a part of all three scenarios.

The network in this case was quite small and the scenarios were quite simple, so no conclusions can be drawn on the applicability of our extraction methods to more complex structures. Nonetheless, we can see that the current reporting format is slightly easier to read than the one proposed in a previous publication (Vlek et al., 2016). A main difference is that subscenarios were previously reported in parentheses within the text of the scenario, as follows:

> Scenario: Bert killed Chris, and [Adam, Bert and Chris were involved in a cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car.] Then Bert moved Chris's body.

In the current report, the scenario only includes the subscenario node 'Adam, Bert and Chris were involved in a cannabis operation', while the detailed subscenario is reported separately, which makes the scenarios easier to read.

6.4.5 Criterion 5: reporting evidential support and scenario quality

The quality of the alternative scenarios did not play a large role in this case, which is why it is hard to evaluate whether it was adequately reported. In any case, it was reported that each scenario was complete and consistent and that each contained the implausible element 'Bert killed Chris'.

The evidential support of each piece of evidence for each scenario was also reported. A noteworthy result is that the DNA evidence is reported to be (only) 'moderate evidence to support' for each scenario. This is because rather than reporting the support of the DNA match for the hypothesis that the DNA belonged to Adam (which we expect to be very high), what is reported is the evidential support related to activity-level hypotheses about how the DNA traces came to be. By reporting this as moderate, we aim to help a judge or jury to better understand the value of forensic evidence such as DNA matching.

A topic for future research is to evaluate experimentally how insightful participants find the proposed reporting of evidential support. Similarly, the usefulness of applying a verbal scale can be subjected to an experiment. We expect that a qualitative report of evidential support is more insightful for a judge or jury than a numeric report, but as was also discussed for criterion 3, some information is lost when translating to a qualitative measure of evidential support.

Part IV: Related work and conclusion

7. Related work

Many researchers have studied reasoning with legal evidence, exploring, for instance, the various features of judicial proof (MacCrimmon and Tillers, 2002) or the various approaches by which legal proof can be analysed (Anderson et al., 2005; Kaptein et al., 2009; Dawid et al., 2011). The research presented in this thesis has combined two topics within the study of legal evidential reasoning: scenario-based reasoning and probabilistic reasoning. In the sections below we discuss related work, divided into related scenario-based approaches (Sections 7.1 to 7.9) and related Bayesian network approaches (Sections 7.10 to 7.19).

7.1 Part 1: Introduction to scenario-based approaches

Scenario-based reasoning in legal cases was first studied from the perspective of legal psychology, dating back to the 1980s. Psychologists found that jurors tend to form scenarios to make sense of the evidence (Bennett and Feldman, 1981; Pennington and Hastie, 1993). This led to the development of theories about how to reason with scenarios, descriptive as well as normative, such as those proposed by Bennett and Feldman (1981), Pennington and Hastie (1993) and Wagenaar et al. (1993). These scenario-based approaches from legal psychology served as a starting point for our work and will be discussed in Sections 7.2 to 7.4 below.

Related to the scenario-based approaches from legal psychology is the theory of explanatory coherence by Thagard (1989), developed from a cognitive science perspective. Though the theory was not specifically aimed at legal applications, Thagard illustrated its application with legal cases. Explanatory coherence is a form of inference to the best explanation (IBE), as are the scenario approaches from legal psychology. In explanatory coherence, alternative explanations are compared on how well they explain the evidence, but also on other properties such as their simplicity. Thagard's explanatory coherence will be discussed in Section 7.5.

Recently, scenario-based approaches have been studied from a computational perspective, leading to the formalisation of some notions from the legal psychologists mentioned above. Bex (2011) has developed a Hybrid Theory for reasoning with scenarios and arguments, in which he formalised several concepts from legal psychology. Bex's Hybrid Theory is discussed in Section 7.6. Keppens and Schafer (2006) also studied scenarios from a computational perspective, leading to the development of a decision support tool and later the combination of this decision support tool with the use of Bayesian networks (Shen et al., 2006). This work will be discussed in Section 7.7. Verheij (2014) has studied the integration of scenarios, arguments and probabilities. His method will be discussed in Section 7.8. Other researchers have worked on scenarios in a legal setting in the form of narratives about the parties involved in a case, such as witnesses, experts or judges (Sileno et al., 2014). This will be discussed in Section 7.9. Each section concludes with a discussion of similarities and differences with our work.

7.2 Bennett and Feldman's study of scenarios in court

Bennett and Feldman (1981) were the first to recognise that scenarios play an important role in the courtroom. They were particularly interested in how people with no specialised training in reasoning with legal evidence were capable of coming to some judgement in a trial, as is done in a jury system. Bennett and Feldman concluded that scenarios can provide a context in which social behaviour can be understood, thereby providing the context needed to judge what has happened. They furthermore argued that scenarios help to store and compare various interpretations of the evidence. This allows people to follow the case and reason about it.

According to Bennett and Feldman, a good scenario is a scenario that has a clearly identifiable central action and a context ('setting') in which this central action makes sense. This way, scenarios serve as mechanisms to determine the relevance of various factors in a case. Bennett and Feldman furthermore pointed out that for a scenario to be believable, it should be consistent, complete and structurally unambiguous.

Bennett and Feldman performed experiments to find out how people reason with scenarios. Participants were asked to assess the truth of various scenarios, some of which were really true and others were not. Some true scenarios were complete and unambiguous, while others were incomplete or at times ambiguous. Similarly, the completeness and ambiguity of the untrue scenarios varied. The results of these experiments showed that complete and unambiguous scenarios were believed more readily by participants than incomplete and/or ambiguous scenarios, regardless of whether they actually happened. This became known as the phenomenon of a good scenario pushing out a true scenario (Anderson et al., 2005).

Discussion The aim of Bennett and Feldman was to study how people reason with legal evidence when they are not using any specific methods. In contrast, our work is aimed at developing a normative method for reasoning with evidence. The work done by Bennett and Feldman serves as an important foundation for our method since it confirms that working with scenarios is an intuitive way to reason with legal evidence. In our method, we aimed to combine this intuitive way of thinking with the normative approach of Bayesian networks.

7.3 Pennington and Hastie's Story Model

Pennington and Hastie (1986, 1988, 1992, 1993) also studied scenario-based reasoning in legal cases from a psychological perspective. Their focus was on understanding how a juror processes the information presented in a trial and uses this to reach a decision. They proposed a theory, called the Story Model, which describes how jurors use scenarios in their decision process. A main claim, which Pennington and Hastie supported with experimental results, is that scenarios not only provide context, but also determine the decision made by a juror.

The Story Model According to Pennington and Hastie, the process from evidence being presented to reaching a verdict consists of three stages: (1) evaluating the evidence through scenario construction, (2) learning the relevant verdict categories and (3) reaching a decision by matching a scenario to a verdict. In these three stages, scenarios are used to organise the information presented at a trial, and are later used to relate to the appropriate verdict to reach a final decision.

In the Story Model, a scenario is a causal chain of events, in which elements are connected by physical causality or intentional causality. Pennington and Hastie furthermore spoke of scenarios being organised in episodes (possibly consisting of subepisodes), referring to previous work on scripts (Schank and Abelson, 1977) and grammars (Rumelhart, 1975). As an example of such an episode, Pennington and Hastie provided an abstract episode schema in which a scenario always consists of initiating events, psychological and physical states, goals, actions and consequences. An illustration of their abstract episode schema is shown in Figure 7.1.

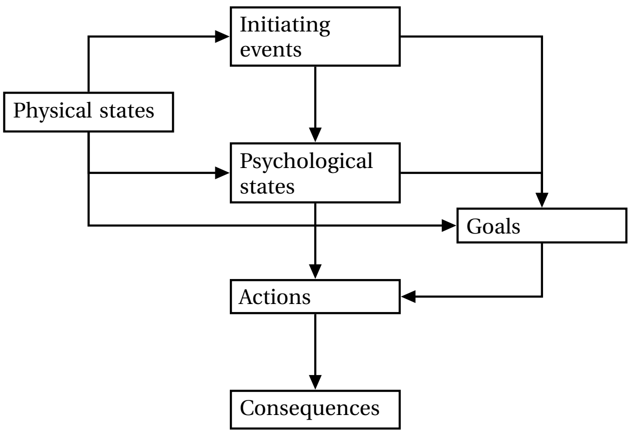


Figure 7.1: An abstract episode schema from Pennington and Hastie (1993), with elements of a scenario connected by causal links.

By processing the trial information via scenario construction, several alternative interpretations of the evidence are given in the form of scenarios. According to the Story Model, a juror uses these alternative interpretations to reach a verdict, matching a scenario to an appropriate verdict category.

Certainty principles The Story Model has four certainty principles that determine which decision jurors will reach and how certain they are of this decision. These certainty principles are coverage, coherence, uniqueness and goodness-of-fit.

Coverage and coherence determine whether a scenario is acceptable as a viable alternative. Pennington and Hastie viewed scenarios as alternative explanations of the evidence and the better the scenario covers the evidence, the more acceptable the scenario is as an explanation of the evidence. When a scenario leaves a lot of evidence unexplained, this results in a low evidential coverage and thus low acceptability. The coherence of a scenario is determined by three factors: completeness, consistency and plausibility. A scenario is complete when it 'has all its parts'. To be consistent, the scenario should not contain internal contradictions. The plausibility of a scenario depends on how well it relates to the juror's knowledge of the world and how things typically work. A scenario may be complete, consistent and plausible to some degree, thereby influencing the degree of acceptability of that scenario.

The other two certainty principles, uniqueness and goodness-of-fit, next determine how certain jurors are of their decision. When multiple scenarios are considered coherent, they are no longer unique. Since this means that multiple alternatives are each acceptable, jurors will be less certain of their final decision. Finally, a juror will need to relate a scenario to a verdict, which relates to the fourth certainty principle, goodness-of-fit. When a scenario does not fit well with a verdict, this decreases the juror's certainty.

Experimental results A central claim of the Story Model is that scenarios are not only used as a context for the final decision; they also determine that decision. Pennington and Hastie supported this claim with experimental results. In one experiment, participants were asked to assess a real trial as if they were the jurors. They were asked to talk out loud while making their decision. It was found that these mock-jurors' mental representation of the evidence was in the form of scenario structures rather than other plausible structures. Furthermore, it was found that people had drawn inferences about what might have happened and would not just consider the evidence as a list of items. Finally, it was found that people who reached different verdicts also had different scenarios in mind, which is consistent with the claim that a scenario determines the verdict.

Discussion Pennington and Hastie developed the Story Model, which is a descriptive model of how people tend to reason with legal evidence, while it also provides a systematic approach to assist people in their reasoning task. Compared to the Story Model, our method takes a formal approach using the framework provided by a Bayesian network. We have formalised Pennington and Hastie's ideas on coherence or goodness of a scenario by probabilistically interpreting the three factors completeness, consistency and plausibility. This way, it becomes possible to point out what happens when a good scenario pushes out a true scenario: a scenario with high plausibility (which is also complete and consistent) is chosen over a scenario that in the end has a higher posterior probability due to the evidence supporting it.

We furthermore used the concept of a scenario scheme (discussed by Pennington and Hastie as an episode scheme) to represent a scenario in a Bayesian network as a cluster of elements. With this cluster-like structure, scenarios are clearly visible in the network and the probabilities calculated in the network include the effect of transfer of evidential support.

By formalising scenarios in terms of probabilities, the selection of the scenario with the highest posterior probability becomes straightforward. This contrasts with Pennington and Hastie's Story Model, in which it is not always clear how to balance the four certainty principles. On the other hand, in our explanation of a network we proposed to not only report the posterior probability of each scenario, but also information about evidential support and scenario quality, returning to a scenario-based approach in which it is left to the judge or jury to balance these various aspects.

7.4 Wagenaar, Van Koppen and Crombag's Anchored Narratives Theory

Building on the notion that people tend to use scenarios for reasoning in criminal trials, Crombag et al. (1992) and Wagenaar et al. (1993) developed their Anchored Narratives Theory. The development of this theory was specifically driven by the observation that there were many 'dubious cases' in court, which might be prevented with the use of a more systematic approach to reasoning with evidence. Whereas the work by Bennett and Feldman and Pennington and Hastie was done in the context of a jury system, the perspective of Wagenaar et al. (1993) was that of a system with only a judge who could use scenarios to rationally compare alternative interpretations of the evidence. The Anchored Narratives Theory is a normative theory for how one should deal with scenarios in a criminal case and Wagenaar et al. illustrated the application of their theory with numerous example cases.

Anchoring A key aspect of the Anchored Narratives Theory is the notion that a scenario should be anchored in our common sense knowledge of the world. This anchoring is important when dealing with evidence, but also when dealing with elements of the scenario in general. For example, when dealing with the evidence one might need to anchor in a common sense rule such as 'an expert witness usually tells the truth'. When anchoring a connection within a scenario (such as 'Bill was angry, so Bill pulled a gun'), one might use the common sense rule 'when people are angry, they may pull a gun'.

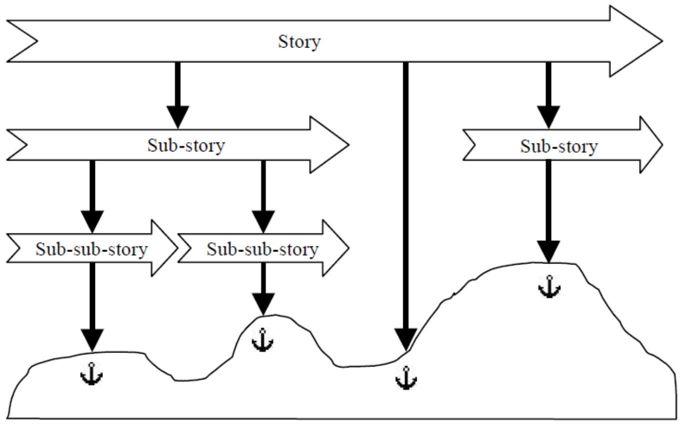


Figure 7.2: An illustration of the Anchored Narrative Theory by Wagenaar et al. (1993), with scenarios (stories) anchored in subscenarios and in common sense knowledge.

The level at which elements of a scenario are anchored may vary. Specifically, a scenario may be anchored by specifying a subscenario that provides more detail about a part of the scenario or about a piece of evidence. Such a subscenario can in turn be anchored in common sense knowledge. An example of such hierarchical anchoring is shown in Figure 7.2.

According to Wagenaar et al., it is crucial that any part of a scenario can be questioned and that one should only believe a scenario when it is sufficiently anchored in so-called safe rules about the world that cannot sensibly be doubted. Although evidence is clearly a way by which a scenario can be anchored, the scenario ultimately needs to be anchored in common sense knowledge (such as: an autopsy report is usually correct) to explain how the evidence (an autopsy report) relates to the rest of the scenario.

Plausibility According to Wagenaar et al., a scenario should be plausible and sufficiently supported by facts. For plausibility or goodness of a scenario, Wagenaar et al. referred to the ideas by Bennett and Feldman, where a good scenario has some central action and a setting in which the central action makes sense. Wagenaar et al.

furthermore observed that even when a scenario seems very implausible, it can still become believable when it is sufficiently anchored in common sense rules.

Universal rules of evidence Wagenaar et al. provided directions as to how their Anchored Narratives Theory should be applied in court by means of a list of ten rules. These so-called universal rules of evidence are as follows:

1. The prosecution must present at least one well-shaped scenario.
2. The prosecution must present a limited set of well-shaped scenarios.
3. Essential components of the scenario must be anchored.
4. Anchors for different components of a scenario should be independent of each other.
5. A judge or jury should give reasons for the decision by specifying a scenario and the accompanying anchoring.
6. A judge or jury's decision should be explained through an articulation of the general beliefs used as anchors.
7. There should be no competing scenario with equally good or better anchoring.
8. There should be no falsifications of the indictment's scenario and nested subscenarios.
9. There should be no anchoring onto obviously false beliefs.
10. The indictment and the verdict should contain the same scenario.

Evidential support In the original Anchored Narratives Theory, the role of evidence is mainly to anchor scenarios in common sense knowledge. In later work Van Koppen (2011) extended the ideas of the Anchored Narratives Theory with a notion of evidential support, to be used when comparing scenarios to verify and falsify which scenario is the true scenario. According to Van Koppen, a piece of evidence supports a scenario more than another scenario when the evidence is more likely to occur in the case of the first scenario than in the case of the second.

Discussion Wagenaar et al. developed their Anchored Narratives Theory as a normative theory, with the anchoring of a scenario as a key feature. Wagenaar et al. were explicitly concerned with the practical application of scenarios in court, which is why they specified a list of ten universal rules of evidence. These universal rules dictate which scenarios should at least be made explicit in a court case and how they should be analysed. In our method, no such list is explicitly specified, but some of the rules by Wagenaar et al. are still incorporated in our method implicitly, although others are beyond our goal of modelling a case since they are concerned with the judge or jury's decision and the verdict (rules 5,6,10). Our method assumes that several alternative scenarios (rules 1 and 2) need to be modelled. Furthermore, the explanatory report contains information concerning the quality and evidential support for a scenario, including the reporting of evidential gaps to find whether essential but implausible components of a scenario are supported (rule 3). Our explanation techniques are not aimed at relating (independent) support to different components of a scenario (rule 4), since evidential support is only reported relative to the scenario as a whole. A combined measure of evidential support is reported for each scenario, allowing for the comparison of evidential support between scenarios (related to rule 7, though anchoring not the same as evidential support). Finally, the use of a Bayesian network should prevent any falsification of the scenario or support in false beliefs within the model (rules 8 and 9).

The key idea of the Anchored Narratives Theory is that miscarriages of justice can be prevented by questioning how well a scenario is anchored in common sense knowledge. In our method, the common sense knowledge is represented in the structure and probabilities of the Bayesian network. Our construction method is designed to capture the plausibility of a scenario in the probability tables of the network. Moreover, we use scenario scheme idioms to capture the common sense knowledge about how scenarios are typically organised.

The definition of evidential support by Van Koppen (2011) is somewhat different from our use of the term evidential support, though the general notion that scenarios should be compared on how well they fit the evidence overlaps. The notion of evidential support proposed by Van Koppen is essentially a likelihood ratio, in which the probability of the evidence given alternative hypotheses is compared. In our method, a different measure of evidential support was proposed, comparing the prior and posterior probability of a scenario given the evidence.

7.5 Thagard's theory of Explanatory Coherence

In Thagard's (1989) theory of explanatory coherence, the best explanation for a set of evidence is selected based on several factors such as coherence, simplicity and whether it is in turn explained by other hypotheses. Thagard's theory of explanatory coherence is an example of inference to the best explanation (IBE) and in that respect similar to scenario-based approaches, which are also a form of IBE (Pardo and Allen, 2008). Thagard (2004) compared his method to Bayesian networks, illustrated with the application to a legal case.

Coherence To formalize what makes a good explanation, Thagard introduced the notion of coherence. Hypotheses and evidence are represented as nodes in a graph with weighted links between them that indicate the coherence between nodes. Coherence links are drawn as regular lines, incoherence links (with negative coherence) are drawn as dashed lines. An example of such a graph is shown in Figure 7.3.

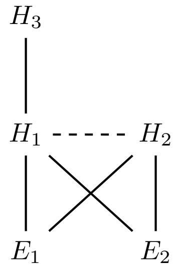


Figure 7.3: An example of a graph representing coherence (regular lines) and incoherence (dashed line) from Thagard (1989) (weight of links not indicated).

Coherence and incoherence links between nodes are symmetric. There are several ways in which two propositions can cohere. For example, when a hypothesis explains another hypothesis, they cohere. A coherence link can also be established when there is an analogy between hypotheses and what they explain, or when alternative hypotheses explain the same evidence. Finally, contradicting hypotheses incohere.

The degree of coherence is represented as the weight of a link, with incoherence as a negative weight (the weight of links is not shown in Figure 7.3). The weight of coherence and incoherence links depends on the number of hypotheses that together explain something: the more hypotheses are needed, the lower the degree of coherence between the hypotheses and the evidence.

ECHO Thagard implemented the principles of his theory of explanatory coherence in a program called ECHO. To determine which explanation should be accepted, each hypothesis has a degree of acceptability, depending on its coherence with other hypotheses and on how well it explains the evidence. This acceptability is expressed as activation in ECHO. Evidence has a fixed degree of acceptability, which is implemented in ECHO by connecting all evidence to a special node that has activation 1. Then the activation is propagated through the structure via the weighted links to find the acceptability (or activation) of each of the alternative hypotheses.

A situation can be modelled in ECHO by entering information about the hypotheses and the evidence, such as which hypotheses explain each other, which are analogous and which contradict each other. Extra parameters can be used to express how certain a piece of evidence is, or how well an explanation explains something. Based on the input of these commands, ECHO draws a diagram of all hypotheses and the evidence with their coherence and incoherence links. Weight is assigned to each link depending on the number of hypotheses that together explain something; ECHO lowers the weight of a link proportional to the number of hypotheses that form an explanation together.

Running ECHO results in the propagation of the activation throughout the diagram. Links with lower weight will diminish the activation and negative links will result in negative activation. After running ECHO, each hypothesis will have some activation assigned to it. Positive activation is interpreted as accepting an hypothesis and the hypothesis with highest activation is the best.

Explanatory coherence compared to Bayesian networks Thagard (2004) illustrated the use of ECHO with a legal case, comparing the theory to a Bayesian network for that same case. Thagard showed how both methods can be used to analyse a legal case and discusses the similarities and differences between the two approaches. Thagard pointed out that a Bayesian network has directed connections between nodes, while the connections representing explanatory coherence are undirected. On a more general level, Thagard's main conclusion is that the theory of explanatory coherence is better as a descriptive model of how a judge or jury reasons, matching the existing ideas on scenario-based reasoning by Pennington and Hastie and others. In contrast, Bayesian networks are normative rather than descriptive. According to Thagard, Bayesian networks are therefore less suitable, particularly due to the requirement that many conditional probabilities need to be assessed, which, according to Thagard, do not always have a clear interpretation.

Discussion Thagard's (1989) theory of explanatory coherence and his implementation in ECHO can be used to select which explanation best explains a set of evidence and is in that respect similar to a scenario-based approach, in which the best scenario is selected. Coherence is an important concept in both approaches and the theory of explanatory coherence provides a formalisation in which the coherence of explanations can be quantified. A difference with the scenario-based approach is that coherent scenarios have an underlying scenario scheme. In our method, we have used these scenario schemes in the form of scenario scheme idioms to help simplify the construction process of a Bayesian network.

There are also similarities between the theory of explanatory coherence and Bayesian networks, most obviously visible in their graphical representation. As claimed by Thagard (2004), the theory of explanatory coherence is close to a judge or juror's intuition. And while Bayesian networks provide a clear mathematical framework for dealing with uncertainty, according to Thagard they are less intuitive, mostly due to the many numbers that need to be specified. With our proposed method, we have attempted to bridge the gap between Bayesian networks and the scenario-based approach, but we recognize that the elicitation of many conditional probabilities is not an easy (nor intuitive) task.

7.6 Bex's Hybrid Theory of Stories and Arguments

Bex has developed a formal hybrid theory for analysing legal cases with arguments and scenarios (Bex, 2011; Bex et al., 2010). Ideas from scenario-based approaches and from argumentation theory were combined, resulting in a method in which scenarios are used as hypotheses about what may have happened and arguments can support scenarios based on the evidence or on common sense knowledge.

Sense-making The research on the hybrid theory was led by the goal to develop a theory that was natural for people to use, rationally well-founded and sufficiently formally specified such that it could lay the foundations for software support systems. Such a software tool could help a fact finder to make sense of a case by visualising his or her reasoning in various ways (Bex et al., 2007). An adaptation of the hybrid theory was used as a basis for a software tool developed by Van den Braak et al. (2007). The resulting software tool was tested in empirical studies with crime analysts, suggesting that the system is indeed natural to use (Van den Braak, 2010).

Evidential arguments and causal scenarios Bex observed that arguments and scenarios provide two distinct ways of reasoning with legal evidence. The argumentative approach can be used to reason from a piece of evidence towards a conclusion, with some evidential rule to justify the conclusion based on the evidence. This approach is typically dialectical in that arguments for and against a conclusion are formed and atomistic since it is focussed on details rather than on the case as a whole. The argumentative approach is very suitable for exposing errors in a chain of reasoning. The scenario-based approach, according to Bex, uses causal reasoning, connecting various events in a scenario with causal generalisations. The scenario-based approach is holistic since it considers the case as a whole, and is a natural way for people with no formal training to think about a case, as was shown by Bennett and Feldman (1981) and Pennington and Hastie (1993).

A combination As argued by Bex, arguments work well when dealing with the evidence, while scenarios can help to form alternative hypotheses about a case. Bex therefore proposed a combination of arguments and scenarios in the form of his Hybrid Theory of Stories and Arguments. This theory is qualified as hybrid because it consists of two parts: an evidential argumentation theory that deals with arguments and a causal explanation theory that deals with scenarios. Arguments and scenarios interact in the theory since arguments can support elements of a scenario and the conclusions of arguments can be explained by a scenario.

Arguments In the Hybrid Theory, arguments are formalised in a logical model of argumentation. An argument can be formed based on the evidential data and a stock of knowledge. This knowledge is formalised as generalisations and can be extended with specific context-dependent reasons for reasoning with evidence, such as 'Witness w says "ϕ" is a prima facie reason for believing ϕ'. The evidence and the stock of knowledge together form the input for forming arguments. Such arguments are finite sequences of elements that are either in the theory, or can be derived from it using generalisations. Arguments can attack each other, as depicted in Figure 7.4. The conclusions drawn from the arguments serve as explananda, things that need to be explained, for the scenarios.

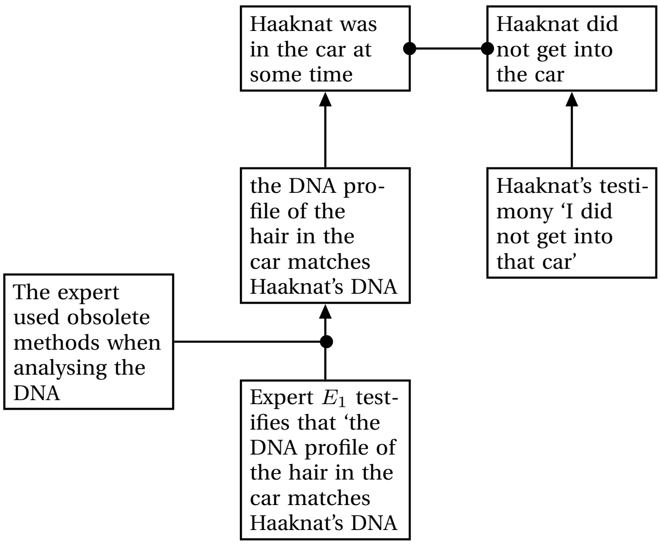


Figure 7.4: An example of attack relations in an argument graph from Bex (2011). Arrows with a closed arrowhead represent evidential generalisations. Arrows with a closed circle at the end represent attacking arguments.

Scenarios In the Hybrid Theory, scenarios are formalised using inference to the best explanation (IBE) as a basis. A scenario consists of events and generalisations connecting these events. The events in a scenario are either hypothetical events that serve as initial elements of the scenario, or events that follow from these initial elements using causal or abstract generalisations. A scenario is, much like the definition of an argument, a finite sequence with initial events and events derived from the previous elements with generalisations. A scenario then forms an explanation for the set of explananda when the scenario is consistent and contains all the explananda as events in the scenario. A graphical representation of a scenario is shown in Figure 7.5.

Coherence To be considered a good explanation, a scenario should be coherent, which was formalised by Bex based on the three factors proposed by Pennington and Hastie (1993): completeness, consistency and plausibility. Arguments can be used to support whether or not a scenario is coherent.

For Bex, internal consistency is incorporated into the definition of an explanation: it is defined as a consistent scenario, one that does not contain internal inconsistencies. The plausibility of a scenario is the extent to which evidential gaps and causal generalisations in the scenario are supported by the stock of knowledge with explicit arguments. Similarly, implausibility is the extent to which elements of the scenario are contradicted by the stock of knowledge, when they are not already contradicted by any of the evidential data.

The completeness of a scenario is defined in terms of scenario schemes. A scenario should both fit and complete a scenario scheme: all elements in the scenario should correspond to something in the scheme, and all elements of the scheme should have a corresponding element in the scenario. Using abstract generalisations, a correspondence between a scenario and a scenario scheme can be established. The scheme to which a scenario corresponds should be plausible: none of the generalisations or components should be contradicted by the evidence or common sense knowledge.

The definition of plausibility in particular requires arguments about scenarios, but arguments can also be about how well a scenario fits or completes a scheme. This shows how the two parts of the hybrid theory are intertwined.

Combining arguments and scenarios A hybrid theory about a case can be used to form arguments as well as scenarios as described above. The conclusions of arguments can occur as elements in the scenario and when they do, the scenario is said to explain these conclusions. An example of a combination of a scenario and arguments is shown in Figure 7.6. This figure shows a scenario at the top, in which horizontal, open-headed arrows connect the elements of the scenario, starting with 'Haaknat is a drug addict who needs money'. Some elements in the scenario are supported by arguments, which are visualised using close-headed vertical arrows. Two pieces of evidence serve as the basis for supporting arguments: 'Expert E_1

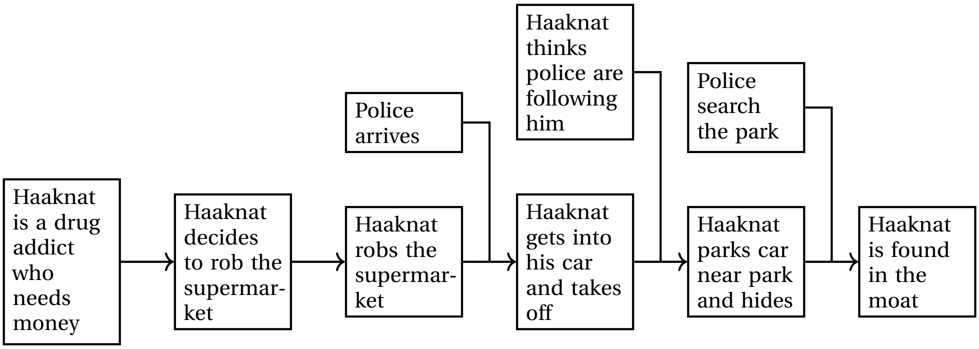


Figure 7.5: An example of a scenario from Bex (2011). Arrows with an open arrowhead represent causal or abstract generalisations.

testifies that the DNA profile of the hair in the car matches Haaknat's DNA' and 'Haaknat's testimony 'I did not get into that car'.' While this evidence is used to support the event 'Haaknat gets into his car and takes off', the scenario serves as an explanation of this event.

Evidential support A piece of evidence supports some event when an argument can be formed with the evidence as a premise and the event as a conclusion. Similarly, evidence can contradict some conclusion when an argument can be formed to attack that conclusion based on the evidence. The evidential support for a scenario then consists of those evidential data that support some event or generalisation in the scenario. Similarly, the evidential contradiction for a scenario is the set of evidential data that contradict some part of the scenario. Events in the scenario that are neither supported nor contradicted by any of the evidential data are evidential gaps.

A dialogue game A hybrid theory as defined above describes a situation with a definitive collection of evidential data, common sense knowledge, etcetera. However, in the investigative process around a crime, such a theory needs to be updated when new evidence is found or new arguments or scenarios are given. In order to capture this continuously changing character, Bex developed a formal dialogue game. The idea of this dialogue game is that several players are making moves in which they can add or retract arguments and scenarios. The players are thereby working on the same hybrid theory, expanding and fine-tuning it up to the point where all players agree on one optimal explanation.

Since the dialogue game depends on the finding of the best explanation, an ordering on explanations is needed. Bex defined an ordering on the explanations

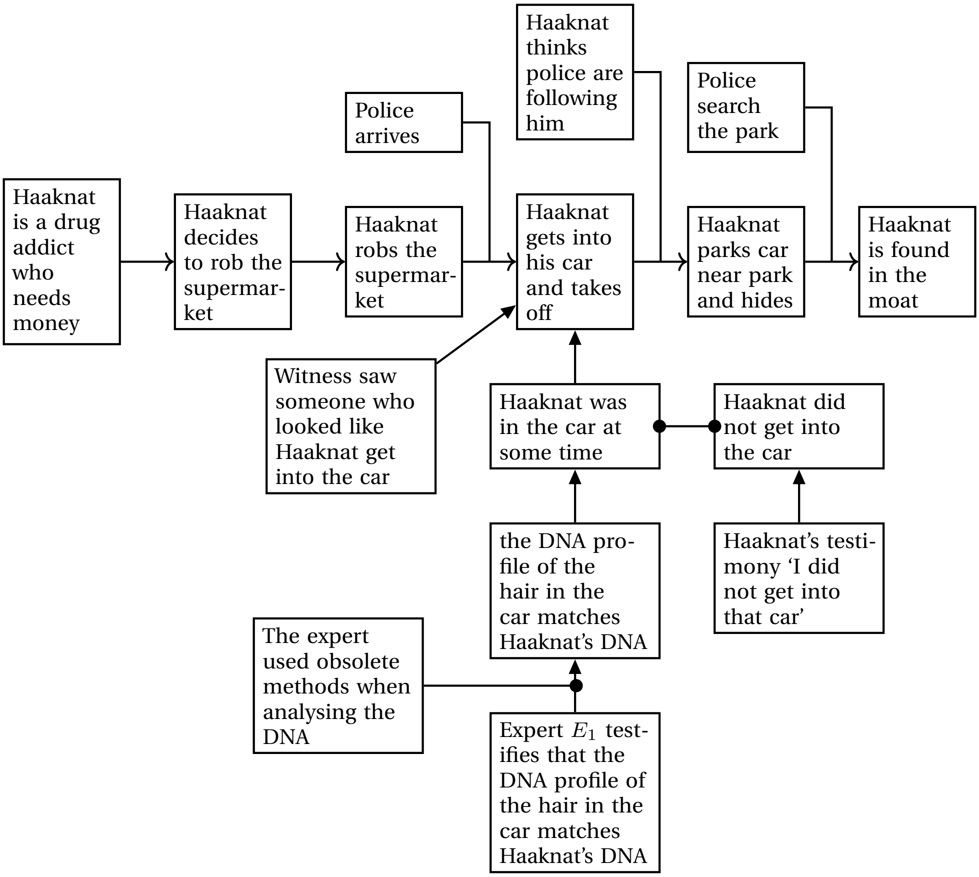


Figure 7.6: An example of scenarios and arguments combined, based on Bex (2011). Arrows with an open arrowhead represent causal or abstract generalisations, arrows with a closed arrowhead represent evidential generalisations. Arrows with a closed circle at the end represent attacking arguments.

using set inclusion. In short, when an explanation explains the same set of events or more (the evidential support is larger or equal) and is contradicted by the same set or less (the evidential contradiction is smaller or equal), it is at least as good or better. When the evidential support and contradiction are equal, two explanations can be compared on plausibility and evidential gaps according to the same principle. As Bex described, this ordering will leave many explanations incomparable. Only when the evidential support of one explanation is really contained in the evidential support of the other, can anything be said about which explanation is better. As soon as explanation S_1 explains event E_1 and S_2 does not, while S_2 explains event E_2 that S_1 does not, there is no conclusive answer. As an alternative, explanations can also be compared on the *number* of events they explain. In that case, an explanation is better when the number of elements in its evidential support is larger and the number of elements in its evidential contradiction is smaller. Bex pointed out that an appropriate choice of ordering may differ per case. The dialogue game will stop when there is some optimal explanation to which all players are committed.

Discussion Bex's Hybrid Theory combines formal argumentation with the holistic approach of scenarios. The theory makes it possible to reason about scenarios with the formal structure of arguments and attacks among arguments. Bex furthermore formalised concepts from the scenario-based approach, such as scenario schemes, coherence and evidential support. Inspired by this formalisation, we have incorporated similar formalisations of notions from the scenario-based approach in our method.

An advantage of our use of Bayesian networks is that they provide a framework in which these narrative concepts can be defined. We gave probabilistic interpretations of inconsistency (when two elements of a scenario together have probability 0 to occur) and plausibility (the prior probability of an element occurring), while our formalisation of completeness is very similar to that of Bex. We also used the scenario scheme idiom to constrain some probabilities in the network to capture the global coherence of a scenario.

A Bayesian network furthermore provides a framework for combining multiple pieces of evidence and comparing their support for alternative scenarios. In contrast, for Bex, comparing the evidential support of alternative scenarios is only clearly defined when the set of evidence for one scenario is a superset of the evidence for the other scenario, in which case the superset clearly lends more evidential support. In all other situations, the evidence needs to be weighed in some way to take into account that one piece of evidence (for example, a DNA match) gives more evidential support than another (such as a witness testimony).

A key aspect of the work by Bex is that arguments for and against modelling choices are made explicit. For example, arguments can support or attack whether a scenario is plausible. In our method the plausibility of a scenario needs to be estimated by a modeller. Such modelling choices are therefore specified in the probability tables and cannot easily be subjected to debate. Another difference is that the method by Bex allows for a dynamical process with the use of a dialogue game, while our method, though using a gradual construction process, assumes that the scenarios that need to be modelled are all available at the start of the construction process.

Bex analysed the Anjum case, which we used as a case study in Chapter 4. Many similarities between the two case studies can be observed, for example in the organisation of specific subscenarios such as how a hit on the head lead to Leo's death, or the drugging of Leo. While the graphical representations differ, much of the underlying structure is the same, in part due to the fact that both methods use scenario schemes as a basis. A clear difference between the two representations is that in Bex's results, arrows between elements of a scenario can be supported with evidence, which is not the case in our Bayesian network. For instance, the toxicology report supports the connection from 'Marjan gave Leo Temazepam' to 'Leo was in a state of impotence', which can be made explicit in Bex's models but not in our method. Nonetheless, the information from the toxicology report is still present in our model since it was used in the probability table of 'Leo was in a state of impotence', but it is somewhat more hidden. Another difference between the two case studies is that our quantitative account, though clearly dependent on the subjective numbers, allows for a conclusion about what probably happened, while the qualitative account from Bex does not allow for drawing a conclusion.

7.7 Keppens et al.'s decision support system

Keppens and Schafer (2006) developed a system that can generate scenarios based on the evidence in a criminal case. Such scenarios can assist the investigation of a crime by suggesting relevant hypotheses and possible evidence that may be expected as a result of these hypotheses.

Scenarios To Keppens and Schafer, a scenario is a description of a combination of situations and events. To compose scenarios, Keppens and Schafer used compositional modelling techniques, in which the first principles of a domain of interest are captured in so-called scenario fragments. These scenario fragments are causal rules which describe relations between states and events. Based on these scenario fragments, the system can compose a scenario space containing the alternative scenarios that explain a given collection of evidence.

The objective of the decision support system is to find all scenarios that explain a set of evidence. In addition, hypotheses are found that are supported by these scenarios, such as, for example, a murder hypothesis and a suicide hypothesis. The system can also identify additional evidence that could be found if a certain scenario were true and the system can identify additional investigative actions by which more evidence might be uncovered.

The procedure The system employs an abductive diagnosis approach. In an initialisation phase, the system starts out with a set of evidence. The next step is a backward chaining phase, in which the system uses the causal rules (in the form of scenario fragments) to find possible causes of the evidence. This leads to a set of hypothetical scenarios. In a forward chaining phase, using the same causal rules, the system can now find possible additional evidence that could be the result of these hypothetical scenarios. Finally, in a consistency phase, any inconsistent combinations of states and events are indicated. Keppens and Schafer provided a formal algorithm by which the system performs these four steps. This way, the system constructs a scenario space with all plausible scenarios and their possible effects. The results are stored in a network that shows causes and effects connected with causal links between nodes.

The decision support system By finding all plausible scenarios and related hypotheses, investigative actions and evidence, Keppens and Schafer aimed to systematise the investigative process. The scenario space is presented to a user as a graph. The user can perform queries to find scenarios that produce certain evidence or support certain hypotheses. According to Keppens and Schafer, this decision support system promotes the consideration of many alternative scenarios and it makes expert domain knowledge available to a less experienced investigator.

Probabilistic scenario spaces In a related paper, Shen et al. (2006) extended the ideas of Keppens and Schafer (2006) to include the use of Bayesian networks. After generating a scenario space as described above, the result is translated to a probabilistic scenario space in the form of a Bayesian network. Variable instances in the scenario space become nodes in the Bayesian network and the terms scenario, hypothesis and evidence now refer to value-assignments of nodes in the network. When a scenario fragment (a causal rule) has a value-assignment of a node A in the antecedent and a value-assignment of a node B in the consequent, an arrow is drawn from A to B in the network. Probabilities of the network are elicited based on the available domain knowledge.

The result is a Bayesian network that represents the scenario space, making it possible to analyse the scenarios in terms of probabilities. Shen et al. emphasised that their system is not meant to evaluate the evidence in court, but rather to assist an investigator in selecting appropriate evidence collection strategies. In particular, Shen et al. proposed to analyse the amount of information that can be obtained by collecting more evidence in terms of entropy. The entropy of the information can be calculated using conventional Bayesian network inference techniques. By aiming to minimise the resulting entropy, an investigator can select an investigative action that would lead to the highest amount of (new) information.

Discussion The decision support system developed by Keppens and Schafer and Shen et al. uses scenarios to help an investigator find hypotheses, additional evidence and investigative actions. A difference between this system and our method is the interpretation of what constitutes a scenario. In the work of Keppens and Schafer and Shen et al., scenarios are composed of states and events like in our method, but there is no further requirement on any properties they should have, while in our method a scenario should be coherent. Furthermore, our method enables the representation and reporting of scenario quality in terms of completeness, consistency and plausibility.

In the work of Shen et al., the resulting Bayesian network represents the states and events of the scenario as separate nodes without further structure to capture them as forming a scenario together. In our method, we specifically capture the coherence properties of a scenario, representing elements of a scenario as clusters in the network, which have the special property that evidence for one element of the scenario influences the other elements of the scenario as well (transfer of evidential support).

Another difference is that Keppens and Schafer and Shen et al. developed their system specifically for an application during the investigation of a crime, while our method is intended for a later phase, in which the relevant alternative scenarios need to be evaluated and compared. Whereas our method assumes that alternative scenarios are available, the system of Keppens and Schafer and Shen et al. can be of assistance in finding these scenarios and collecting more evidence. Their system is thus especially helpful as a tool that can help prevent tunnel vision by suggesting alternative scenarios that have not been considered during the investigation, but it is not meant to draw conclusions about which scenario is best.

7.8 Verheij's method using scenarios, arguments and probabilities

Verheij (2012, 2014) has aimed to combine scenarios, arguments and probabilities in an integrated approach. This way, according to Verheij, the adversarial setting of arguments, the global perspective of scenarios and the gradual uncertainty of probabilities can be combined.

Scenarios For Verheij, a scenario is a conjunction of individual events, with no particular order. A scenario serves as a complex hypothesis about what may have happened and can also lead to additional expectations about what evidence may be found. Several alternative scenarios are compared in a theory about the case.

A combined approach In the method of Verheij, scenarios form the hypotheses in a case, while arguments can be formed to support a scenario with evidence. The varying strengths of arguments can be expressed numerically as conditional probabilities, but only need to be specified when these numbers are available. A scenario can also have a prior probability expressing its coherence, enabling calculations to find its posterior probability given the evidence. The relation between scenarios, arguments and probabilities is depicted in Figure 7.7.

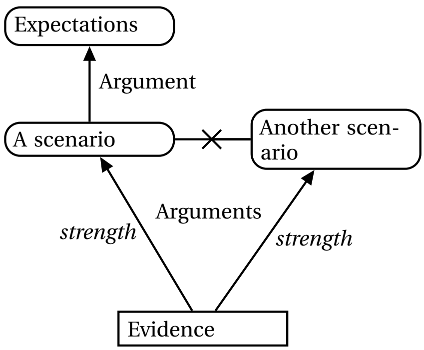


Figure 7.7: The connection between scenarios, arguments and probabilities as depicted in Verheij (2014). Arrows represent supporting arguments and the crossed-out line represents that the scenarios are competing hypotheses.

Theory construction In the method of Verheij, a theory about a case is gradually constructed. Argument strength is made explicit as a probability when numbers are available; otherwise, no numbers need to be specified. This results in a partially specified probability function which forms a theory about the case. When all but one scenario have probability 0 in this probability function, the case is decided. When multiple alternative scenarios are still considered possible, the case needs to be investigated further.

Discussion Verheij used scenarios as complex hypotheses, which can be supported with arguments. Probabilities come into play only when they are available. The use of scenarios as complex hypotheses is a similarity with our work. A difference is that the method of Verheij does not consider the quality of a scenario. In particular, the notion of completeness of a scenario is not present in Verheij, although one might argue that an incomplete scenario would receive a lower prior probability.

The method by Verheij uses standard probability theory. While our method requires a full Bayesian network to be specified to enable any calculations, the method by Verheij only requires partial probabilistic information. This is a clear advantage in a situation in which numbers are typically not all available. On the other hand, the method by Verheij might not have sufficient information to reach a conclusion on a case when few or no numbers are available. In contrast, our method will be able to compute a posterior probability for each scenario, but requires many numbers to do so.

7.9 Sileno et al.'s representation of legal scenarios

Sileno et al. (2014) have studied how legal scenarios can be represented in an agent-based model. In this representation, agents can be, for example, the parties involved in a case, witnesses, experts or judges. The interactions between these agents are interpreted as scenarios and are represented in petri-nets. This approach lends itself well to the representation of social aspects of a case, including for example goals and motives.

Discussion The work by Sileno et al. is concerned with representing the setting in the courtroom in the form of scenarios. This differs from the purpose of our method, which is aimed at fact-finding. Another difference is that Sileno et al. aimed to represent scenarios, not to compare and draw conclusions from them. In a Bayesian network constructed with our method, the posterior probabilities of scenarios can be compared and a conclusion can be drawn about which scenario is most probable given the evidence.

7.10 Part 2: Introduction to Bayesian network approaches

For forensic experts, Bayesian networks provide a useful tool for analysing legal evidence (Dawid and Evett, 1997; Taroni et al., 2006). Bayesian networks are currently typically applied to analyse a specific part of a case that has to do with some forensic evidence. For example, a Bayesian network can be used to analyse how a DNA match relates to the hypothesis that the suspect is the donor of a trace (Evett and Weir, 1998), but also to analyse how two crimes are related via the same modus operandi (De Zoete et al., 2015). Often, pre-fabricated Bayesian network structures can be designed to analyse typical situations, such as, for instance, the analysis of evidence related to fire incidents (Biedermann et al., 2005). Laskey and Mahoney (1997) studied how such fragments can be used together to construct a more complex network, applied to a military situation. Modelling a legal case in its entirety is not a common application of a Bayesian network, though a notable exception is the work by Kadane and Schum (1996), which will be discussed below in Section 7.11.

To aid a modeller with the construction of a Bayesian network for a complex case, construction methods have been proposed. Hepler et al. (2004) first proposed the use of Object-Oriented Bayesian networks (OOBNs, developed by Koller and Pfeffer

(1997)) in a legal setting. With these OOBNs, it becomes possible to incrementally construct a Bayesian network top-down, using modules or fragments to gradually construct a network. This will be discussed in Section 7.12 below. Hepler et al. also introduced the notion that a network can be constructed by using ready-made network fragments which can be applied throughout various cases. This led to the development of a list of legal idioms by Fenton et al. (2013). These idioms will be discussed in Section 7.13. The concept of reusable network fragment was also the basis of the Hypothesis Management Framework developed by Van Gosliga and van de Voorde (2008), which is generally applicable and not specifically intended for the legal field. The Hypothesis Management Framework will be discussed in Section 7.14. Finally, Timmer et al. (2015b) aimed to capture arguments in a Bayesian network by using network fragments for argument schemes, which will be discussed in Section 7.15.

An overview of explanation methods was compiled by Lacave and Díez (2002), which will be discussed in Section 7.16. Henrion and Druzdzel (1990) developed an explanation method in which scenarios are extracted from a network as explanations of the evidence. This work will be discussed in Section 7.17. Others have used arguments to explain Bayesian networks, such as Keppens (2012), who extracted argument diagrams, and Timmer et al. (2015a), who produced an intermediate framework called a support graph, to extract arguments from a Bayesian network. These argument explanations will be discussed in Sections 7.18 and 7.19, respectively.

7.11 Kadane and Schum's analysis of the Sacco and Vanzetti case

Kadane and Schum (1996) performed an extensive analysis of the famous Sacco and Vanzetti case as a study of 'complex inference on large masses of evidence'. In 1921, Nicola Sacco and Bartolomeo Vanzetti were convicted for first-degree murder because they supposedly robbed and killed two men transporting a payroll. A debate arose because many people found that Sacco and Vanzetti had had an unfair trial. Kadane and Schum re-analyse the case, focussing on the construction of inference graphs using Wigmore charts (Wigmore, 1931), combined with a Bayesian analysis of the evidence.

Using the ideas of Wigmore, Kadane and Schum formulated hypotheses and subhypotheses for the Sacco and Vanzetti case and investigated how these relate to the evidence. This resulted in large diagrams similar to Wigmore's original charts, but modernised, based on the work by Anderson and Twining (1991). These inference diagrams represent the chains of reasoning from evidence to conclusion.

Kadane and Schum, and later Schum (2001), observed that such inference diagrams are always directed acyclic graphs (DAGs), as are Bayesian networks. However, as they remarked, Bayesian network algorithms rely on independence properties in the network, which inference diagrams or Wigmore charts do not necessarily satisfy. Nonetheless, according to Schum, a Bayesian analysis can be performed on a Wigmore chart "as long as the chart has been made to represent non-independence correctly" (Schum, 2001). To this end, Kadane and Schum adapted their original Wigmorean inference networks to enable a Bayesian analysis in Bayesian network software tool ERGO.

Discussion The work by Kadane and Schum is exceptional in its attempt to analyse a complex case in its entirety using probabilities. This results in complex inference diagrams. Bayesian networks are only shown for parts of the case, but it is clear that a Bayesian network for the entire Sacco and Vanzetti case would be very complex indeed. With our method, we aimed to simplify and systematise the construction and understanding of a Bayesian network for an entire case. Though this still results in complex networks, the structures are annotated to enable an explanation of a network. In this thesis we have also presented two case studies that were concerned with entire cases. Though our case studies were not as extensive as the study in Kadane and Schum, these case studies showed the advantages of a method for constructing and explaining a Bayesian network for a complex case.

7.12 Hepler, Dawid and Leucari's object oriented Bayesian networks for legal evidence

Hepler et al. (2004) observed that a probabilistic analysis of a complex legal case requires a systematic approach for the handling of different varieties of evidence. They proposed a combination of Wigmore charts as an argumentative tool and object-oriented Bayesian networks (OOBNs) as a probabilistic framework. This led to a method in which a Bayesian network can be gradually constructed using the concepts from Wigmore charting. The method is illustrated with the Sacco and Vanzetti case from Kadane and Schum (1996).

Wigmore charts According to Hepler et al., Wigmore's (1931) method can assist in describing and organising the evidence in a case. Such a Wigmore chart is especially helpful when dealing with large masses of evidence, since the chart helps to keep track of relations between evidence and conclusions drawn from that evidence. In particular, Wigmore was interested in how the evidence relates to the ultimate probandum, the ultimate hypothesis about whether the suspect committed the crime. This can be found by first relating the ultimate probandum to penultimate probanda, which are intermediate hypotheses that need to be evaluated to establish whether the ultimate hypothesis holds. These can in turn be related to additional intermediate hypotheses, until they can be connected to the evidence. Once a

Wigmore chart is constructed, it can be used to follow the reasoning processes in a case.

Bayesian networks Hepler et al. stated that a Bayesian network provides a framework for modelling the relationships between variables in a case, for weighing the evidence and drawing inferences from that evidence. However, Hepler et al. observed that there is no systematic approach to handling the mass of evidence when constructing a Bayesian network.

Object-oriented Bayesian networks Hepler et al. proposed the use of object-oriented Bayesian networks (OOBNs) for gradually constructing a Bayesian network for a legal case. Using the technique of OOBNs, developed by Koller and Pfeffer (1997), a network is built top-down, first indicating relations between high-level objects. These objects can either consist of single nodes, or of more complex structures of other objects. A network can then be constructed gradually, by first specifying the relations between objects and later specifying these objects. Hepler et al. observed that this technique provides the tools for handling the large masses of evidence in a legal case.

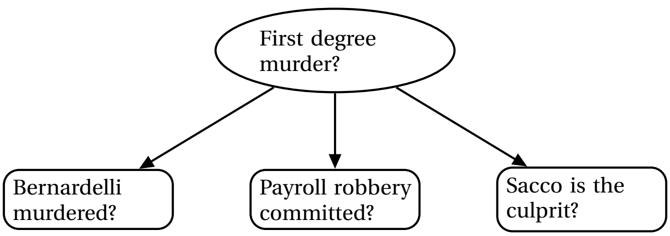


Figure 7.8: A top-level object-oriented network structure for the murder charge against Sacco from Hepler et al. (2004).

Combining Wigmore charts and OOBNs Hepler et al. proposed to combine the rigorous handling of evidence from Wigmore charts with the qualitative and quantitative analysis that is possible in a Bayesian network. Wigmore's notion of dividing up the ultimate probandum in penultimate probanda can be translated to a Bayesian network with the use of OOBNs: the ultimate probandum is drawn as a single node, while the penultimate probanda are drawn as modules or objects that later need to be specified further. This method differs from the approach by Kadane and Schum (1996), who also used arguments and Bayesian networks, but did not have a systematic approach to the gradual construction of a network structure.

Hepler et al. illustrated their method with the famous Sacco and Vanzetti case from Kadane and Schum (1996), as shown in Figure 7.8. To find whether Sacco is guilty of first-degree murder in the slaying of Alessando Berardelli during a robbery (the ultimate probandum, drawn as a single node in the network), one must find whether Berardelli died of gunshot wounds, whether he was in possession of payroll at the time he was shot and whether it was Sacco who intentionally fired shots that took Berardelli's life during the robbery of payroll. These penultimate probanda are modelled as objects in the network that require further specification in a later stage.

Network fragments The ideas proposed by Hepler et al. enable a gradual construction of a Bayesian network, adding more details incrementally. Hepler et al. furthermore proposed a number of network fragments to model often recurring patterns of evidence, such as structures for modelling explaining away and corroboration of evidence. These substructures can be used as 'off-the-shelf' modules in the construction process. This idea was further developed by Fenton et al. (2013), as will be discussed in the next subsection.

Discussion Borrowing ideas from object-oriented Bayesian networks and Wigmore charts, Hepler et al. proposed a method that can be used to gradually construct a Bayesian network. The result is similar to our notion of unfolding a scenario, thereby gradually constructing a Bayesian network with more and more detail. Both provide a way to start the construction of a Bayesian network on a high level and then to gradually work out the details on a lower level. However, there are two ways in which our method differs from the work by Hepler et al.

Firstly, in our approach, at any step of the construction process the resulting structure is a Bayesian network graph with nodes that can all be interpreted as nodes in a Bayesian network. This is because a subscenario node is itself an element of a scenario and whether or not that element is unfolded to a subscenario does not matter for the interpretation of the node. This is not the case for object-oriented Bayesian networks, since these include special objects that still need to be replaced with Bayesian network fragments (or modules) (Hepler et al., 2004). As remarked by Hepler et al., this also means that the construction of such object-oriented Bayesian networks requires dedicated software in which such special objects can be represented and later replaced with fragments. Whereas a network constructed with our method may be visually easier to understand with special software in which subscenarios can be shown folded or unfolded[1], any Bayesian network tool can be used for the construction. However, the explanation techniques in our method, including the annotation of a network during the construction phase, would require dedicated software.

[1] Software-tool GeNIe has the possibility of representing submodels, but these can only be opened in a separate view, not unfolded as part of the larger network.

Secondly, our construction method provides a list of steps guiding the construction of a Bayesian network, including a list of three questions for determining whether an element of a scenario requires unfolding. This differs from the work of Hepler et al., in which no such procedure exists. In particular, Hepler et al. included no directions on how to determine when a Bayesian network contains enough detail.

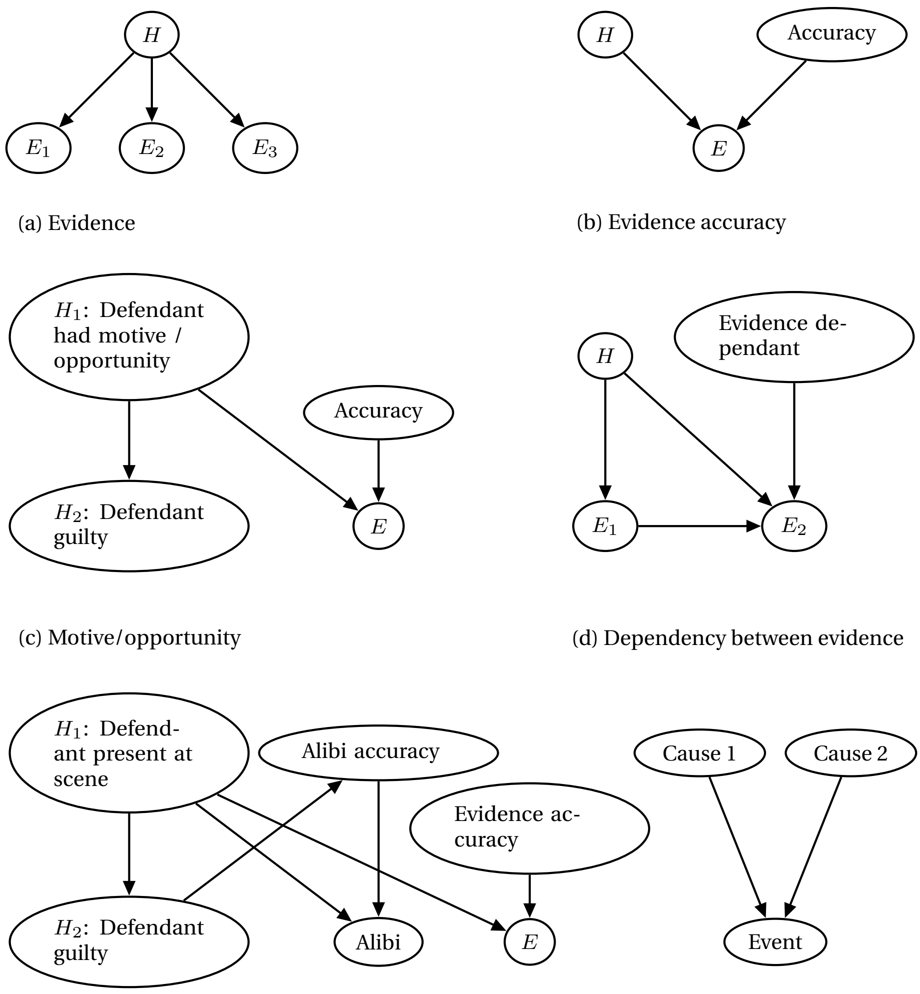
7.13 Fenton, Neil and Lagnado's legal idioms

A motivation for the development of legal idioms by Lagnado et al. (2012) and Fenton et al. (2013) was the observation that no systematic repeatable method for modelling reasoning with legal evidence in Bayesian networks was available. Lagnado et al. and Fenton et al. proposed such a systematic method in which they model the legal case using legal idioms as basic building blocks. This idea of often recurring substructures was inspired by work by Hepler et al. (2004), which was discussed in the previous section. The legal idioms are meant to complement and extend the work by Hepler et al., further developing the use of recurring substructures.

Causal structures Underlying the idioms proposed by Fenton et al. is the idea that a Bayesian network should be constructed such that it represents causal connections. All idioms are therefore designed to represent recurrent patterns in a causal way. Fenton et al. remarked that the resulting network is then consistent with the predominant approach to reasoning with legal evidence: the scenario-based approach from legal psychology as described by Pennington and Hastie and others (see Sections 7.2, 7.3 and 7.4).

The idioms The list of idioms as developed by Fenton et al. is shown in Figure 7.9. The most basic of their idioms is the evidence idiom (Figure 7.9(a)). This deals with a very basic case in which evidence is considered and a hypothesis stating what could have caused this evidence is evaluated. This idiom does not distinguish between evidence that supports the prosecution and evidence that supports the defence in the sense that the structure of the Bayesian network remains exactly the same for both.

Fenton et al. used this elementary idiom as a basis for more elaborate idioms. For example, the evidence accuracy idiom (Figure 7.9(b)) also takes into account that the evidence may have been the result of an inaccuracy. This is modelled with an additional node, the accuracy of the evidence. This node is used to specify the prior probability that the evidence is somehow inaccurate; it is also possible to model different nodes for different types of inaccuracy such as objectivity, veracity and competence for a witness testimony.



(a) Evidence

(b) Evidence accuracy

(c) Motive/opportunity

(d) Dependency between evidence

(e) Alibi evidence

(f) Explaining away

Figure 7.9: Legal idioms from Fenton et al. (2013).

In their idiom for motive and opportunity (Figure 7.9(c)), Fenton et al. extended the evidence accuracy idiom with a node that precedes both the hypothesis and the evidence. This node can either describe an opportunity (the defendant was present at the scene) or a motive (the defendant needed money). In both cases, this should be something that caused both the hypothesis (the defendant is guilty) and the evidence (such as fingerprints).

The idiom for dependency between evidence (Figure 7.9(d)) is specifically suitable for modelling multiple pieces of evidence that are interrelated. In the case that part of the evidence is supplied by the defendant himself, as is the case with an alibi, it can be modelled with the alibi evidence idiom (Figure 7.9(e)). When the hypothesis that the defendant is guilty becomes more likely, this influences the accuracy of the evidence (or, in other words, the credibility of the alibi). Therefore, there is a link from the hypothesis to the accuracy of the evidence.

And finally, when there are two possible causes for one piece of evidence, reasoning about one of them can lead to more or less certainty about the other. This is a common phenomenon in Bayesian networks, called intercausal interaction or (in specific cases) explaining away (Jensen and Nielsen, 2007). Fenton et al. modelled this as shown in Figure 7.9(e), the explaining away idiom.

Discussion The legal idioms proposed by Fenton et al. can be used as ready-made network fragments when constructing a Bayesian network. Our idioms have a similar purpose and are meant to complement and extend the list of idioms as proposed by Fenton et al. As specified in our construction method from Chapter 3, the legal idioms from Fenton et al. are incorporated in our 5-step procedure for constructing a Bayesian network. After the scenarios have been represented, these legal idioms should be applied to relate the evidence to these scenarios.

The legal idioms from Fenton et al. as well as the recurrent patterns from Hepler et al. focus on representing the relations between the evidence and a hypothesis of interest. In particular, Fenton et al. assumed that there is only one ultimate hypothesis of interest with no alternative hypotheses other than the negation of the ultimate hypothesis. In contrast, our idioms can be used to represent alternative scenarios that serve as competing hypotheses about what may have happened, which are not necessarily jointly exhaustive. Moreover, our idioms were developed to represent complex hypotheses in the form of scenarios, whereas the ultimate hypothesis in Fenton et al. consists of a single event.

7.14 Van Gosliga and Van de Voorde's Hypothesis Management Framework

Other research that uses a method for the gradual construction of a Bayesian network was done by Van Gosliga and van de Voorde (2008), who developed the Hypothesis Management Framework (HMF) as a method for constructing a Bayesian network. These HMFs take a modular approach, enabling the specification of details about a case without losing perspective on the case as a whole. By including probabilities explicitly as nodes in the network, multiple experts can work on a single network simultaneously. The incremental aspect of the construction process thus refers to various steps in the construction being delegated to various experts. For example, one expert could work on (part) of the network structure, while another subsequently fills in the corresponding numbers. HMFs are particularly well suited for frequent changes in the network, making them particularly interesting for the legal field, where insights might change during the investigative process.

Discussion The HMF's of Van Gosliga and van de Voorde can potentially be used for the incremental construction of a Bayesian network, though no step-by-step procedure for their use was given. Our method provides a procedure for constructing a Bayesian network in five steps, using scenarios as a guide. These scenarios form the context that can help to find which variables are relevant to a case, while the gradual unfolding of a scenario helps to keep track of the construction process.

7.15 Timmer et al.'s idiom approach for capturing argument schemes

Timmer et al. (2015b) have worked on the representation of argument schemes in a Bayesian network. Other related work by Timmer et al. concerning the explanation of a Bayesian network with arguments is discussed in Section 7.18.

An argumentation scheme captures a pattern of argumentation that may occur in a legal case; Timmer et al. focussed on the specific argument scheme concerned with testimony evidence, in which the evidence 'Witness says "X is true"' supports the conclusion 'X is true'. Such an argument has exceptions, for example, when a witness lies. Such exceptions are taken into account in the argument scheme as so-called critical questions. When dealing with testimony evidence, the critical questions are: (1) Veracity: is the witness sincere?, (2) Objectivity: did the witness's memory function properly? and (3) Observational sensitivity: did the witness's senses function properly?

To represent an argument scheme as an idiom, the evidence (Witness says "X is true") is represented as a separate node, as is the conclusion (X is true). To represent the critical questions, Timmer et al. used ideas from signal filtering, such that the conclusion "X is true" is connected to the the evidence via a sequence of consecutive weaker claims, namely 'Witness perceived X' and 'Witness believes X'. Timmer et al. furthermore proposed to explicitly include the critical questions as parents of these consecutive claims, as shown in Figure 7.10. The advantage of representing the critical questions explicitly is that they are now clearly visible in the network and each have their own prior probability which can be subject of debate. Timmer et al. concluded that a similar approach may be applicable to other argument schemes.

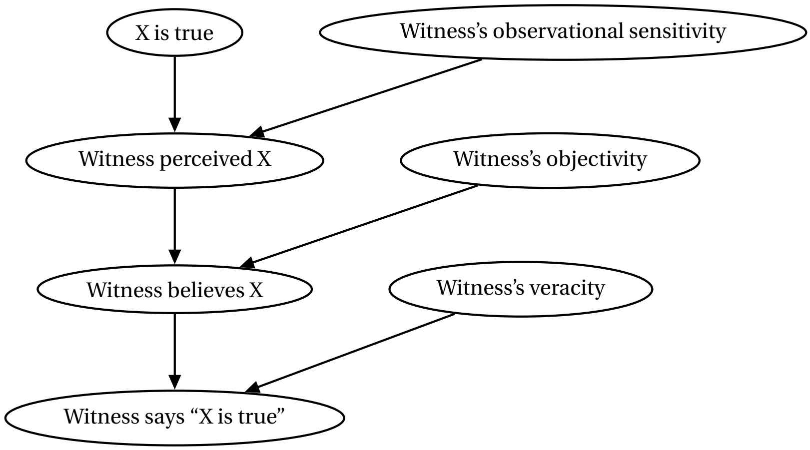


Figure 7.10: Timmer et al.'s representation of the argument scheme for testimony evidence.

Discussion The representation of argument schemes in a Bayesian network as proposed by Timmer et al. is related to our representation of scenario schemes in a Bayesian network. Both are aimed at simplifying the construction process of a Bayesian network, while simultaneously capturing argumentation or scenarios, respectively. While argument schemes have a standard format (an argument from evidence to conclusion, accompanied with a list of critical questions), scenario schemes do not have such a standard format. As a result, Timmer et al. were concerned with the specific structure to capture an argument scheme, while our work has been focussed on the overall construction of a Bayesian network representing various alternative scenarios.

Another noteworthy difference between our work and that of Timmer et al. is the different roles of arguments and scenarios in the construction of a Bayesian network: whereas argumentation is concerned with relating evidence to some hypothesis about what may have happened, scenarios are used to form complex hypotheses. An interesting topic for further research is the possible integration of the two approaches, in which scenario-based idioms can be used to represent the hypotheses in a Bayesian network, while argument-based idioms can be used to relate these hypotheses to the evidence.

7.16 Explanation methods for Bayesian networks

An overview of explanation methods for Bayesian networks can be found in Lacave and Díez (2002). They categorised the various types of explanation methods depending on their content. According to Lacave and Díez, three types of content can and should be explained: the evidence, the model and the reasoning.

Explaining the evidence According to Lacave and Díez, an explanation of the evidence consists of a configuration of unobserved variables. The goal is then to find the most probable explanation (MPE), which is the assignment w of values to the unobserved variables that leads to the highest a posteriori probability given the evidence e: $\Pr(w|e)$ is maximised. Sometimes the goal is to find an assignment to some subset of unobserved variables (described as partial abduction or maximum a posteriori (MAP) estimation) and sometimes the goal is to find the best value assignment for all unobserved variables (total abduction). Algorithms have been developed to find the MPE or the k most probable explanations given a set of evidence. Bayesian network software (including, for instance, GeNIe and SamIam) often include a tool for finding the MPE given a set of evidence.

Explaining the model One way of explaining a given model is simply to show the associated graph in a visually appealing way (Lacave and Díez, 2002). Especially when a Bayesian network graph grows larger, it can become difficult to understand the model, which is why some software tools (such as GeNIe) include the option of organising the graph using submodels, which can be shown separately. Other explanations of the model include colouring arrows to indicate the underlying conditional probabilities (Lacave et al., 2001; Koiter, 2006) and verbal explanations as proposed by Henrion and Druzdzel (1990). The latter is somewhat related to our use of scenarios and will be discussed in Section 7.17 below.

Explaining the reasoning Methods for explaining the reasoning in a Bayesian network typically focus on a single hypothesis, generating explanations for this hypothesis (Lacave and Díez, 2002). Explanation methods then produce verbal or graphical descriptions of how one variable influences another. Most relevant to the legal field is work that focusses on reporting the most relevant nodes that may affect the posterior probability of a certain hypothesis. For example, Suermondt (1992) developed a method for identifying the evidence that influences a certain hypothesis and paths in the network through which this evidence flows. The chains of reasoning can then be presented graphically or verbally. Others have worked on the (verbal) reporting of reasoning chains within the network, see, for instance, the work by Henrion and Druzdzel (1990) in Section 7.17. Similar goals underlie the work by Timmer et al. (2015a), which will be discussed in Section 7.18 below.

Discussion The main focus of our explanation method from Chapter 5 is the explanation of the model. By reporting which scenarios were modelled and how their quality (in terms of completeness, consistency and plausibility) was represented, the content of the model and the result of modelling choices are conveyed. As mentioned by Lacave and Díez, such an explanation of the model can also aid the construction and refinement of the model by providing feedback.

Regarding the explanation of evidence in the model, our method differs from typical methods as described by Lacave and Díez. Often, explanation methods explain the evidence by finding the most probable explanation (MPE). In our method, there is no need to find a most probable explanation by searching for the best assignment to nodes in the network, since the alternative scenarios already serve as alternative explanations of which at most one can be true. The most probable scenario can be read from the proposed report from Chapter 5, since this report includes the posterior probability of each scenario node.

Finally, our method provides insight into the results obtained by the network and the reasoning behind them, by reporting how strongly each piece of evidence supports each scenario. Chains of reasoning are not explicitly reported in our method, but by learning how the pieces of evidence influence the posterior probability of each scenario node, a judge or jury can understand why one scenario is more probable than another. By including evidential gaps in the report, some insight is gained into what hypothetical reasoning might lead to other results.

7.17 Henrion and Druzdzel's explanation method with scenarios

The use of scenarios for explaining a Bayesian network has not been studied much. An exception are the verbal explanations proposed by Henrion and Druzdzel (1990) and Druzdzel (1996), which use scenarios to explain the model as well as the reasoning in the network.

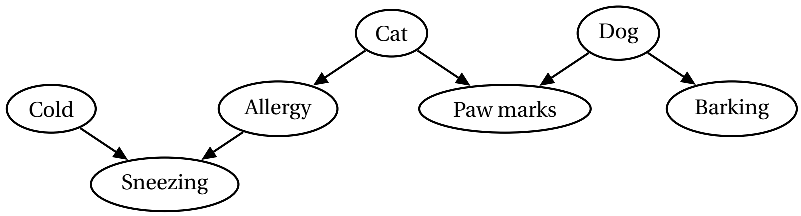


Figure 7.11: The graph of a Bayesian network example from Henrion and Druzdzel (1990).

Verbal explanations of the model To explain the content of a Bayesian network, Henrion and Druzdzel proposed to produce verbal explanations of the causal influences that are present in the network. To convey the strength of an influence, they proposed to report a qualitative verbal description of the underlying probability, which they argued is easier for people to understand than a number. A table for translating probabilities to adjectives and adverbs is provided in Henrion and Druzdzel (1990). For their example network (the graph of which is shown in Figure 7.11), some examples of explanations are:

- 'Cold is very unlikely ($\Pr = 0.08$)'
- 'Cat is unlikely ($\Pr = 0.1$)'
- 'Dog is unlikely ($\Pr = 0.1$)'
- 'Cat commonly ($\Pr = 0.8$) causes allergy.'
- 'Dog as often as not ($\Pr = 0.5$) causes barking.'

Scenarios as explanations of the reasoning Henrion and Druzdzel also proposed to use scenarios to explain the reasoning in a network. To Henrion and Druzdzel, a scenario is a configuration of unobserved nodes in the network, much like the most probable explanation (MPE) described in Section 7.16. They argued that people can understand scenarios better when they are presented as coherent causal stories. To this end, the variables in a scenario are ordered such that effects follow their causes and causes are conjoined when appropriate. As an example, in Figure 7.11 consider `cold` as an hypothesis of interest, given the observations of `sneezing` and `paw marks`. Scenarios as presented by Henrion and Druzdzel now at least include a statement about the hypothesis of interest (`cold`) and includes other hypotheses when they are needed to explain the observations. Some example scenarios from Henrion and Druzdzel are the following:

- 'Cold and no cat hence no allergy and dog.'
- 'No cold and cat causing allergy and dog.'

Discussion Henrion and Druzdzel proposed the use of verbal explanations of inferences in a network and scenarios as configurations of nodes explaining the evidence. Similar to Henrion and Druzdzel, who put nodes in a causal order, we proposed to extract a scenario in text form from the network by ordering the elements according to the graphical structure of the network.

A difference with our use of scenarios is that for Henrion and Druzdzel, a scenario has no further properties beyond being a collection of states or events. For example, with their methods, the following value-assignment would be considered a scenario in our burglary example from Figure 5.7: Sylvia and Tom had debts = `F`, a window was already broken = `F`, Sylvia and Tom climbed through the window = `T` and Tom stole a laptop = T. In our method this would not be considered a good scenario, since some of the elements are false and thus essentially absent from the scenario, making it incomplete.

Another difference is that our method is intended to connect to a scenario-based approach. The reported scenarios and their quality and supporting evidence is meant to provide a judge or jury with sufficient information to consider the case in terms of a scenario-based approach such as those from legal psychology. In contrast, the explanations proposed by Henrion and Druzdzel use scenarios as the end result of a Bayesian network analysis, with no further narrative interpretation.

7.18 Timmer et al.'s explanation of Bayesian networks with arguments

Timmer et al. (2013, 2014, 2015c,d,a) have also worked explaining Bayesian networks for legal evidence, generating arguments based on the content of a given network. Specifically, they proposed to extract a so-called support graph from a Bayesian network, which shows the possible inferential steps in a Bayesian network and can be used to form arguments.

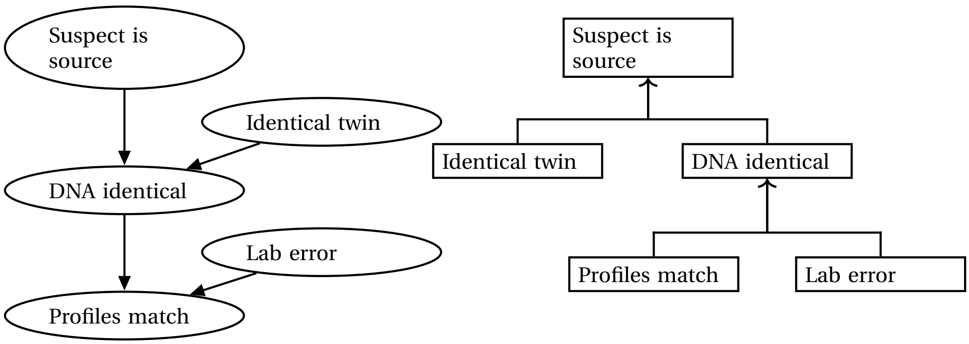


Figure 7.12: An example Bayesian network (left) and the resulting support graph (right) from Timmer et al. (2015a).

The idea of a support graph is that it shows for a specific variable of interest which variables in the Bayesian network possibly provide support for that variable and which variables in turn provide support for these variables, and so on. For a variable V, the variables in the network that possibly provide support are called support factors and they can be found using the Markov blanket for that variable: parents of V, children of V and parents of children of V. Once these support factors are found, each of them in turn have a set of support factors given by their Markov blanket. Some caution is needed when working with a head-to-head connection;

while the support factors for V may include a child U of V as well as U's parent W, the set of support factors for child node U should not include parent W, and W is therefore considered a forbidden node. This construction is needed to exclude arguments such as 'it rained (W), therefore the grass was wet (U), which means that the sprinklers must have been on (V)'.

A support graph can be constructed incrementally by finding support factors for the variable of interest, and support factors for these support factors, taking into account forbidden nodes. Such a support graph is shown in Figure 7.12. A support graph can be viewed as an intermediate format between a Bayesian network and an argument graph, which can help to construct arguments based on the information in the Bayesian network. A support graph is distinctly different from an argument graph because the nodes in the support graph represent variables rather than propositions; by assigning values to these variables, an argument can be formed.

Discussion Timmer et al. proposed the concept of support graphs, which can be used to generate arguments based on a Bayesian network. The argument extraction method proposed by Timmer et al. applies to any Bayesian network, whereas our explanation techniques require that a network was made with our construction method. Although applicability to any network is clearly an advantage, with our combined construction and explanation method we can be sure that the network is constructed such that meaningful scenarios can be extracted from it. This is not necessarily the case for the method of Timmer et al.

As mentioned in Section 7.15, the roles of arguments and scenarios differ since arguments are concerned with relating the hypotheses to the evidence while scenarios are concerned with complex hypotheses. This also means that arguments and scenarios can be useful in different ways for explaining a Bayesian network. Our method focussed mainly on explaining which scenarios are in the network and how they are represented. Our explanation also included the reporting of evidential support, but does not explicate how or why a piece of evidence supports a hypothesis. The latter can be reported using the method by Timmer et al., so an interesting topic for future work would be to combine the two explanation methods to optimally use the advantages of scenarios as well as arguments.

7.19 Keppens' argument extraction methods

Related work by Keppens (2011, 2012) aims to extract arguments from a Bayesian network in the form of argument diagrams. Keppens argued that argument diagrams and Bayesian networks are closely related and can be used in complementary ways. In particular, according to Keppens, an argument diagram is easier to understand than a Bayesian network and it is specifically suitable to test inferences for potential flaws. With these features, an argument diagram can be useful to test and revise a Bayesian network, which can help to improve the construction of a Bayesian network.

Keppens proposed an algorithm for extracting an argument diagram from a Bayesian network. This algorithm can be applied to extract an argument diagram for a specific situation, in which some variable of interest and some set of evidence are known. The resulting argument diagram shows the various arguments that can be made based on the evidence to support or refute the hypothesis of interest.

Discussion Keppens proposed to produce argument diagrams based on a given Bayesian network. As discussed in the previous section, arguments and scenarios can play different roles when explaining a Bayesian network. In particular, argument diagrams may be unsuitable for explaining a large and complex Bayesian network, since the resulting argument diagrams will also be complex. A main advantage of our method is that scenarios are very suitable for understanding a large and complex network, since scenarios can help to make sense of the case as a whole.

7.20 Conclusion

In this chapter, scenario-based approaches and Bayesian network approaches to reasoning with legal evidence were discussed and related to the research presented in this thesis. Our research was led by the goal to combine scenarios and probabilities. Research from legal psychology showed that scenarios are close to a judge or jury's intuition, which led to the development of scenario-based methods, some descriptive and some normative. In recent work by Bex (2011), scenarios and arguments have been combined in a normative theory, which was used as the basis for a software support system developed by Van den Braak (2010). In our method we aimed to combine the intuitive nature of scenarios with the probabilistic framework of Bayesian networks.

A combination of these two approaches has not been studied much, with the exception of the decision support system by Shen et al. (2006) The work by Shen et al. differs from our research in its application: while our method can be used to evaluate a fixed collection of scenarios in a criminal trial, the method by Shen et al. is to be used during the investigative process, assisting an investigator by generating possible scenarios. The method by Shen et al. is explicitly not intended to evaluate which scenario is best.

Our research questions were specifically concerned with the construction and understanding of Bayesian networks, and how scenarios might be of help. Specifically, we proposed a method by which a Bayesian network could be constructed with a 5-step procedure using four narrative idioms, scenario schemes, the unfolding of scenarios and a probabilistic interpretation of scenario quality. Related work on constructing Bayesian networks is often based on the notion of reusable network fragments, such as the legal idioms by Fenton et al. (2013). These network fragments or idioms are similar to our use of idioms. They differ in the scope of the proposed fragments, which typically focus on the local structure of a Bayesian network. With our narrative idioms we extend the existing idioms with a global perspective that can help to construct the network for a case as a whole.

Another key aspect of our construction method is the concept of unfolding a scenario, by which a network could be gradually constructed. Hepler et al. (2004) proposed a gradual construction method using object-oriented Bayesian networks. This gradual construction is similar to our use of unfolding scenarios, but the work by Hepler et al. differs from ours since it only leads to a Bayesian network graph once all details have been specified. In contrast, our construction method can be used to specify a Bayesian network at various levels of detail and at each level the result is a Bayesian network graph. Another difference is that we included a 5-step construction procedure, which helps to determine how and when more detail should be added.

Our second research question was concerned with explaining a Bayesian network. As an extension to our construction method we proposed explanation techniques by which scenarios, scenario quality and evidential support could be reported. Related work on explaining Bayesian networks is sparse. Mostly, an explanation of a Bayesian network is concerned with the most probable explanation (MPE), which is the configuration of unobserved nodes with the highest posterior probability given the evidence. The explanation method proposed by Henrion and Druzdzel (1990) employs scenarios to report something similar to an MPE. However, their use of scenarios differs from our interpretation of what constitutes a scenario, since they are not concerned with coherence or scenario quality. In our explanation method, scenario quality is explicitly reported.

Other explanation methods are often aimed at explaining the chain of reasoning from evidence to conclusion, such as the explanation methods using arguments as proposed by Timmer et al. (2015c) and Keppens (2012). Whereas our scenario-based explanation method is advantageous because it employs the global perspective provided by scenarios, it does not include an explanation of such chains of reasoning from evidence to conclusion. It can be beneficial to extend our proposed reporting format with such chains of reasoning in the form of arguments.

8. Conclusion

In this thesis we have investigated a combination of a probabilistic approach and a scenario-based approach to reasoning with legal evidence in criminal cases, resulting in a method for constructing and explaining Bayesian networks with scenarios. Scenarios can be represented in a Bayesian network, while they also function as a guideline for the construction and explanation of the network. In this chapter we evaluate our results in light of the two research questions (Section 8.1), followed by a discussion of wider implications (Section 8.2). Finally, we present suggestions for future research (Section 8.3) and end with some concluding remarks.

8.1 Results

Two main research questions motivated this thesis. The first question was concerned with the construction of a Bayesian network for legal evidence and how ideas from a scenario-based approach might aid the construction process. The second question was concerned with the explanation of a Bayesian network for legal evidence and how it might be assisted with ideas from a scenario-based approach. In the sections below, the results concerning each research question are discussed.

8.1.1 Research question 1: construction

Research question 1. How can tools from scenario-based reasoning aid the construction of a Bayesian network modelling legal evidence in a criminal case?

As part of our construction method in Part II of this thesis, we proposed the use of four narrative idioms, which provide the graphical structure needed to represent several alternative scenarios in a Bayesian network graph. By representing several alternative scenarios with these idioms, the construction of a Bayesian network is assisted since these scenarios provide descriptions of the domain that needs to be represented, thereby suggesting which variables are relevant for the modelling of a legal case. Moreover, because of the use of scenario scheme idioms, ready-made structures are available to model typical crimes, such as a burglary, a murder, and so forth. By selecting a matching scenario scheme idiom for a given scenario, a basic layout of nodes and connections between nodes is available. These idioms thus use the narrative notion of a scenario scheme to provide structure to the construction process.

In the Anjum case study (Chapter 4), we found that the use of scenario scheme idioms indeed provides structure to the modelling process. In particular, even when scenario scheme idioms need to be developed 'on the fly' for lack of a database with scenario scheme idioms, it makes the construction of the network less ad hoc than without the use of scenario schemes, since the modeller is prompted to think about the underlying structure of a scenario.

The scenario scheme idiom also allows for the representation of scenario quality and more specifically completeness. Definitions of consistency and plausibility were also provided. Though the representation of scenario quality is not directly of assistance to a modeller during the construction of a Bayesian network, thinking about the notions of completeness, consistency and in particular plausibility can aid the assessment of probabilities. In the Anjum case study, we found that while the quality of a scenario can be captured adequately, the assessment of plausibility is not straightforward. This suggests that the narrative notion of plausibility does not sufficiently simplify the assessment of probabilities.

Using the four narrative idioms and the interpretation of scenario quality, a network can be incrementally constructed by unfolding a scenario into more detail when required. A 5-step construction procedure was presented for the incremental construction of a Bayesian network representing several alternative scenarios. This construction procedure uses ideas from scenario-based reasoning by employing the global perspective of scenarios and the fact that (sub)scenarios can be unfolded to various levels of detail. Our Anjum case study showed that the 5-step construction procedure and in particular the unfolding of scenarios is helpful to keep an overview when working with a large and complex network. However, the case study also showed that the last step, that of specifying the probabilities, is disproportionally large and that the method could benefit from more assistance to a modeller in this step.

Contributions In sum, the contributions of this thesis with respect to the first research question are as follows:

1. Four narrative idioms for representing scenarios in a Bayesian network;

2. Scenario scheme idioms as ready-made structures based on the concept of scenario schemes;

3. A probabilistic interpretation of scenario quality in terms of completeness, consistency and plausibility;

4. A 5-step construction procedure for incrementally constructing a Bayesian network representing alternative scenarios, their quality and the evidence;

5. A case study to evaluate the proposed construction method.

8.1.2 Research question 2: explanation

Research question 2. How can tools from scenario-based reasoning aid the understanding of a Bayesian network modelling legal evidence in a criminal case?

While our construction method may result in a rather large and complex network, the graph is annotated and structured using the narrative idioms. Moreover, with the use of scenario scheme idioms, the representation of a scenario in the network is closely related to the original scenario, since nodes in the network correspond to propositions in the scenario and arrows correspond to connections. This way, the narrative notion of scenario schemes is used to aid the explanation of a Bayesian network. In particular, the construction with scenario scheme idioms allowed for the extraction of scenarios from the network. These scenarios in text form are used in a report about the contents of the Bayesian network, including the evidence as instantiated in the network.

As was found in the Nijmegen case study (Chapter 6), in an explanation of the network these scenarios can convey the content of the network and specifically the various alternative explanations of the evidence. The case study also showed that an explanation in terms of scenarios conveys the content of the network at a global level, but does communicate the local structure of the Bayesian network graph.

In addition to a textual report of the scenarios in the network, we proposed a reporting format which includes the scenarios as they were modelled in the network, as well as information about their quality and the evidential support. The explanatory report can be used by a judge or jury to understand a Bayesian network in terms of a scenario-based approach. The report contains all the ingredients needed for a scenario-based analysis of the case, but enhanced with probabilistic information. For each scenario, its posterior probability given the evidence is reported, as well as the degrees of evidential support of each piece of evidence and the combined evidential support for all evidence. Finally, (probabilistic) information about the quality of a scenario is reported. This proposed reporting format uses ideas from the scenario-based approach, namely that to decide on a case one should consider the alternative scenarios, their quality and the evidence supporting these scenarios. The Nijmegen case study revealed that these three ingredients indeed give insight into the reasoning behind a case, since the report for the Nijmegen case matched well with the findings of the actual appeal case.

We furthermore observed that scenario quality, particularly plausibility, affects how evidential support changes the probability of the scenario, thus providing insight into why a scenario might have a higher posterior probability than another scenario. We observed that, when an element of a scenario is initially very implausible, evidential support for that element has a stronger effect on the probability of the whole scenario than evidential support for a more plausible element of the scenario. Similarly, an implausible element of a scenario that remains unsupported, called an evidential gap, can be of great importance for understanding why a scenario is not more probable. In our proposed report, we specifically point out implausible elements with and without evidence supporting them. The notion of scenario quality is thereby used to aid the explanation of a Bayesian network.

Contributions In sum, the contributions of this thesis regarding the second research question are as follows:

1. A procedure for extracting scenarios in text form from a Bayesian network built using scenario schemes during the construction;
2. A reporting format for reporting scenario concepts from a Bayesian network in terms of scenarios, scenario quality and evidential support, enhanced with probabilistic information;
3. Evaluation of the proposed explanatory report by means of a case study.

8.2 Putting the results into context

The research questions of this thesis were formulated in the context of a higher-level problem, namely that of combining the various existing approaches to reasoning with legal evidence. In particular, the research in this thesis has been concerned with a combination of a scenario-based approach with a probabilistic approach, aiming to inherit strengths of both approaches. In this section, we discuss how well our method addresses this higher level problem.

Table 1.1 in the introduction (page 10) showed various characteristics of each of the three approaches: the argumentative approach, the scenario-based approach and the probabilistic approach. In Table 8.1 an extended version of this table is shown, now elaborated with how our method incorporates the strengths from the scenario-based approach and the probabilistic approach. Each of these strengths will be discussed below, followed by a discussion of some limitations inherited from the two approaches.

Strengths An adversarial setting can be found in the scenario-based approach in the form of alternative scenarios. By always considering multiple alternative scenarios, a scenario-based approach can help to reduce the risk of tunnel vision (Wagenaar et al., 1993). In the spirit of this approach, our proposed method explicitly requires that multiple alternative scenarios are represented in a network. As a result of the narrative idioms, each scenario is clearly represented as a separate structure in the network and separate scenario nodes identify the alternatives in the network. With the merged scenarios idiom, the alternative scenarios are represented in a single network, which enables the contrasting and comparing of alternatives as is

	Argu-ments	Scenarios	Probabilities
Adversarial setting	✓	(✓) Multiple alternative scenarios are represented in a single network Scenarios are presented as alternative hypotheses in the network's explanation	-
Global perspective	-	✓ Scenarios are represented as clusters of events The explanatory report places evidence in their scenario context	-
Degrees of uncertainty	(✓)	-	✓ Plausibility of a scenario is represented in degrees Evidential support reported in degrees
Closeness to intuition	✓	✓ Scenarios assist the construction of a Bayesian network Scenarios are used to explain a Bayesian network	-
Standard formalisation	(✓)	-	✓ Bayesian networks have standard semantics

Table 8.1: Characteristics of the three approaches to reasoning with legal evidence. A check mark ✓ denotes that an approach has this characteristic, while a check mark in parentheses (✓) denotes that the approach has this characteristic only to a limited extent. A hyphen - denotes that the characteristic is absent in this approach.

customary in the scenario-based approach. In the explanatory report, the scenarios are again presented as alternative hypotheses in a case.

The global perspective of a scenario provides a context to the evidence (Pennington and Hastie, 1993). In our method, this global perspective is taken into account by representing scenarios as clusters in a network. With the probabilities, the scenarios are modelled such that they behave as coherent clusters. In particular, the transfer of evidential support is captured in this way, meaning that when an element of a scenario is supported by evidence, the other elements of that scenario also receive evidential support as a result of coherence properties. The global perspective of scenarios is also advantageous in the explanation of the network, since this can help a judge or jury to make sense of the evidence. Our method therefore reports evidential support for scenarios (instead of evidential support for separate events).

Degrees of uncertainty are a strength of the use of probabilities, and Bayesian networks provide the formal framework needed to deal with degrees of uncertainty (Taroni et al., 2006). In our method we have captured scenarios in a Bayesian network. Specifically, the plausibility of a scenario can be expressed in our method in the prior probabilities of the scenario node and the plausibility of elements of the scenario can be expressed in the probabilities of element nodes. As in any probabilistic approach, the uncertainty of evidence can also be captured in our method. The proposed explanatory report includes evidential support reported in (qualitative) degrees. If a judge or jury were to use this report to assess the case in a scenario-based approach, these degrees of evidential support are a clear advantage: while a traditional scenario-based approach has no formally specified way of comparing evidential support for alternative scenarios (unless the evidence of one scenario is clearly a superset of the other), our method enables the weighing of evidence for alternative scenarios, including the weighing of the combined mass of evidence for each scenario.

The scenario-based approach has been shown to relate to a judge or jury's intuition (Pennington and Hastie, 1993). The research in this thesis was motivated by the question how this could be used as an advantage when constructing and explaining a Bayesian network. During the construction of a network, scenarios serve as a guideline for selecting which variables are relevant and the unfolding of scenarios is used to gradually construct a Bayesian network for a case. During the explanation of a network, scenarios provide the context for the evidence and the scenario-based approach provides a framework which specifies which features of a Bayesian network should be reported: the scenarios, their quality and the evidential support.

The standard formalism of Bayesian networks provides a framework in which our method could be specified. In addition, we have provided probabilistic interpretations of key concepts in the scenario-based approach: coherence, evidential support (in the narrative sense) and scenario quality in terms of completeness, consistency and plausibility. But while the mathematical underpinning of Bayesian networks means that there is agreement on how calculations in the network should be performed, modelling a Bayesian network for a real-world problem is a whole different exercise. First steps towards a systematic modelling approach for legal cases were taken by Hepler et al. (2004) and Fenton et al. (2013), who proposed idioms to be used as network fragments. With the use of scenarios, our goal was to systematise the construction of a Bayesian network, with a focus on the graphical structure needed to represent a complex legal case as a whole.

Limitations While we employed strengths of each approach as much as possible in our proposed combination of scenarios and probabilities, our research showed that there are some limitations as well. The most pronounced limitation of our method relates to the use of Bayesian networks, which require that many numbers need to be specified before calculations can be performed (Jensen and Nielsen, 2007). This is a well-known issue with the use of Bayesian networks and in our method we have focussed on the construction of the graph, referring to existing elicitation techniques for finding the numbers. Such elicitation methods are available (see, e.g., (Renooij, 2001)), but it is not yet clear how well they perform in applications to the legal field, or in combination with our proposed method.

Another issue is the informality of the scenario approach. Scenarios can be rather open constructions, in which much is understood from context. To represent scenarios in a Bayesian network, they need to be made explicit. By doing so, our formalisation of the scenario-based approach in a probabilistic framework can partially resolve some of the limitations that come with the informality. For instance, as described above in relation to degrees of uncertainty, the probabilistic approach was used to formalise the notion of evidential support. While a pure scenario-based approach has no strict way of comparing evidential support, in our method there is now a mathematically correct way to compare the evidential support for one scenario to the evidential support for another. Nonetheless, some informality from the scenario-based approach remains in our method. For instance, the scenario-based approach does not specify how scenario quality and evidential support can be weighed to reach a conclusion. It is then up to a judge or jury to consider the trade-off between a good scenario and a well-supported scenario. In our proposed report, scenario quality and evidential support are included, but similar to a scenario-based approach, it is up to a judge or jury to determine how scenario quality and evidential support are interpreted.

8.3 Future work

Elicitation of probabilities From the fact that we use Bayesian networks as a probabilistic framework, it follows that many probabilities need to be specified, since a Bayesian network cannot be used to perform calculations unless all probability tables have been specified. Our focus in this thesis has been on the construction of the graph of a Bayesian network rather than on the probabilities.

Although our method specifies some probabilities as part of the four narrative idioms, many probabilities remain to be elicited during the construction of a network. For instance, in the Anjum case study the resulting network consisted of 61 nodes, for which each probability table required between 1 and 8 numbers to be assessed (excluding the numbers that are already specified by the idioms or that follow since probabilities need to sum to 1). Thinking about the notion of plausibility instead of probability could help the assessment of probabilities, but shifting the task from assessing probabilities to assessing plausibility might not be much of an improvement.

In future research it would be interesting to study under what circumstances a Bayesian network is useful and usable in a legal case. Additionally, it might be of interest to investigate the use of a sensitivity analysis in combination with our method. The constraints on the probabilities as dictated by the idioms could make it easier to perform such a sensitivity analysis, since the number of probabilities that need to be varied in such an analysis is reduced.

Further evaluation Another topic for future research is to further evaluate the proposed method. In this thesis we have presented two case studies as an evaluation of the method, which were intended to test whether the method was useful when applied to a real case. During the four years of research we have informally had students perform case studies with our method as part of their bachelor projects, mostly focussing on the construction of a Bayesian network. Three cases were modelled in these projects, namely the Wamel murders (Israëls, 2006), two mysterious shootings (Crombag et al., 2009) and the mud murder (Scherrenburg et al., 2013). From these case studies, our bachelor students concluded that the method was helpful to keep an overview of the case while constructing the method. A noteworthy anecdote is that one student initially had trouble constructing a Bayesian network, only to realise that he had forgotten to use our construction method.

In future research, it would be informative to further evaluate the method and in particular to investigate with a group of realistic participants, such as professional judges, how insightful they find the proposed explanation method and the qualitative report of evidential support as part of that explanation method. Such an evaluation might also show whether chains of reasoning from evidence to conclusion are clearly lacking from the explanation, since this is something that is not made explicit in our scenario-based explanation.

Theory construction Currently, our construction method is focussed on representing a collection of scenarios in a Bayesian network. It is assumed that these scenarios are available, for instance, as a result of a police investigation. In further research, it may be of interest to explore a more dynamic construction procedure, in which a Bayesian network can be used to represent scenarios that need to be further investigated by the police, elaborating the network with new finding as the investigation proceeds. Several elements of our construction method can prove to be helpful during the investigative process: scenario scheme idioms could be helpful in generating hypothetical scenarios around some event and the process of unfolding could help to keep an overview of the main scenario while investigating details of a case. Finally, the explanation techniques proposed in this thesis could be helpful by providing feedback about the construction of a network in a dynamic setting. However, using a Bayesian network in a dynamic setting is not straightforward, since a small change in the graph may lead to many new probabilities that need to be assessed.

Integration with arguments Finally, a main topic of this thesis was the integration of a scenario-based approach with a probabilistic approach to evidential reasoning. An evident topic for future research is the further integration with the argumentative approach. Such future research could not only be directed towards an integrated perspective, but also towards an improved method for constructing and explaining a Bayesian network.

Throughout this thesis, we have argued that scenarios provide a global perspective on a case and are therefore helpful for representing complex hypotheses in a Bayesian network, guiding the incremental construction of a network. As for the explanation of a network, scenarios can be used to explain which complex hypotheses were modelled in a Bayesian network, including their scenario quality and evidential support. As we remarked in our chapter on related work (Chapter 7), the role of scenarios differs from that of arguments, since arguments are often more focussed on relating the evidence to a hypothesis, possibly via intermediate hypotheses. Scenarios and arguments can therefore function in complementary ways when constructing and explaining a Bayesian network.

To further explore how scenarios and arguments can support the construction and explanation of Bayesian networks, future research could investigate a combination of narrative and argumentative idioms during the construction phase. In a related project, Timmer et al. (2015b) has worked on the development of argumentative idioms. In a combination, our narrative idioms could be used to represent complex hypotheses, while Timmer et al.'s argumentative idioms could connect these hypotheses to the evidence. Timmer et al. (2015a) have also worked on the explanation of Bayesian networks with arguments, generating arguments for or against a certain hypothesis of interest based on a given network. Our proposed explanatory report could benefit from the addition of such arguments to show the chains of inference from the evidence to each scenario.

8.4 Concluding remarks

A main motivation for this thesis was the observation that the analysis of statistical evidence in court can be problematic, in particular when it needs to be incorporated in the context of a case as a whole. While forensic experts are advocating the use of Bayesian methods (Aitken, 2012; Berger and Aben, 2010b,c,a), such probabilistic approaches are not typically insightful for a judge or jury. In this thesis, we studied a combination of scenarios and probabilities, with a particular focus on how scenarios could aid the construction and explanation of Bayesian networks.

The result is a method in which scenarios guide the construction of a Bayesian network, while representing scenarios, their quality and the evidence in the network. As an explanation of the network, a report is compiled with scenarios extracted from the network, scenario quality and evidential support. Case studies showed that the proposed method indeed assists the construction and the explanation of the network.

In our research, we have used Bayesian network as a formal framework. While it is clear how probabilistic calculations should be performed in a Bayesian network, much debate is possible on how to correctly model the domain of interest, such as a legal case. With our method, we aimed to further systematise the construction of a Bayesian network in legal applications. While our method was not focussed on the assessment of probabilities but rather on finding the graphical structure of a Bayesian network, we have encountered the issue that many numbers need to be made explicit in a Bayesian network. Because of this, further research is needed to find how Bayesian networks might be used in court, before any method (including the one here) is ready for applications in actual cases.

List of publications

Journal articles

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014a). Building Bayesian networks for legal evidence with narratives: A case study evaluation. *Artificial Intelligence and Law*, 22(4):375 – 421.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law*. In press.

Verheij, B., Bex, F., Timmer, S., Vlek, C., Meyer, J.-J., Renooij, S., and Prakken, H. (2016). Arguments, scenarios and probabilities: Connections between three normative frameworks for evidential reasoning. *Law, Probability & Risk*, 15(1):35–70.

Peer-reviewed conference proceedings

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013a). Modeling crime scenarios in a Bayesian network. In *Proceedings of the 14th International Conference on Artificial Intelligence and Law*, pages 150 – 159, New York. ACM Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013b). Representing and evaluating legal narratives with subscenarios in a Bayesian network. In Finlayson, M., Fisseni, B., Löwe, B., and Meister, J., editors, *2013 Workshop on Computational Models of Narrative*, pages 315–332, Saarbrücken/Wadern, Germany. Schloss Dagstuhl.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013c). Unfolding crime scenarios with variations: A method for building Bayesian networks for legal narratives. In Ashley, K., editor, *Legal Knowledge and Information Systems: JURIX 2013: The Twenty-Sixth Annual Conference*, pages 145 – 154, Amsterdam. IOS Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014b). Extracting scenarios from a Bayesian network as explanations for legal evidence. In Hoekstra, R., editor, *The 27th International Conference on Legal Knowledge and Information Systems (JURIX) 2014*, pages 103–112, Amsterdam. IOS Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2015a). Constructing and understanding Bayesian networks for legal evidence with scenario schemes. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law*, pages 128– 137, New York. ACM Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2015b). Representing the quality of crime scenarios in a Bayesian network. In Rotolo, A., editor, *Legal Knowledge and Information Systems. JURIX 2015: The Twenty-eighth Annual Conference*, pages 131 – 140, Amsterdam. IOS Press.

Nederlandse samenvatting

In een strafzaak krijgt een rechter of jury een collectie bewijs ter overweging. Op basis van een rationele redenering moet men vervolgens een conclusie trekken over wat er mogelijk gebeurd is rondom de misdaad, alvorens een oordeel over schuld of onschuld kan worden geveld. Met de komst van moderne forensische technieken wordt een rechter of jury steeds vaker geconfronteerd met een gekwantificeerde onzekerheid over een bewijsstuk. Dit is bijvoorbeeld het geval bij DNA-bewijs: de rapportage van een match wordt doorgaans vergezeld van een zogenaamde *random match*-kans, die aangeeft hoe groot de kans is dat deze DNA-match gevonden zou zijn als de verdachte niet de bron is van het spoor. Aan de rechter of jury is vervolgens de niet eenvoudige taak om het bewijs in de zaak als geheel in beschouwing te nemen, inclusief kwalitatief bewijs (zoals getuigenverklaringen) en kwantitatief bewijs (zoals DNA-bewijs).

Om te redeneren met juridisch bewijs heeft een rechter of jury de beschikking over verschillende redeervormen, doorgaans ingedeeld aan de hand van de onderliggende benaderingen: de argumentatiebenadering, de scenariobenadering en de kansbenadering (Anderson et al., 2005; Kaptein et al., 2009; Dawid et al., 2011). Elke benadering heeft zijn eigen voor- en nadelen. Hoewel de kansbenadering zeer geschikt is voor het rekenen met onzekerheden, zoals de bovengenoemde *random match*-kans, zijn de argumentatiebenadering en de scenariobenadering voor een rechter of jury intuïtiever in het gebruik, en beter geschikt voor het overwegen van de zaak als geheel. Een combinatie van of vertaling tussen verschillende benaderingen is het onderwerp van onderzoek (Verheij et al., 2016). Hoewel het gebruik van kansen in de rechtszaal ter discussie staat, richten forensische experts zich steeds meer op het gebruik van een kansbenadering, in het bijzonder met behulp van Bayesiaanse netwerken, waardoor het des te meer relevant wordt om te bestuderen hoe zo'n kansbenadering inzichtelijk kan worden gemaakt voor een rechter of jury. In dit proefschrift bestuderen we daarom de mogelijkheid om verschillende benaderingen te combineren, met het achterliggende idee dat zo'n combinatie de voordelen van de verschillende benaderingen kan samenvoegen.

In eerder onderzoek is er reeds aandacht geweest voor de combinatie van argumenten en scenario's in het werk van Bex (2011). De combinatie van argumenten en kansen is bestudeerd door Timmer et al. (2015a), en Verheij (2014) onderzocht de combinatie van argumenten, scenario's en kansen. Een overzicht van deze verschillende onderzoeksprojecten en hun connecties is te vinden in Verheij et al. (2016). Dit proefschrift heeft als onderwerp de combinatie van scenario's en kansen. De probleemstelling is dan ook als volgt:

Probleemstelling. Hoe kunnen een scenariobenadering en een kansbenadering gecombineerd worden in het redeneren met juridisch bewijs in strafzaken, op zo'n manier dat voordelen van beide benaderingen optimaal gebruikt worden?

Zoals gezegd is de kansbenadering geschikt voor het redeneren over de onzekerheid die hoort bij statistisch bewijs. Momenteel wordt de kansbenadering vooral gebruikt op kleine schaal, bijvoorbeeld voor de analyse van een onderdeel van een zaak. Een scenariobenadering is geschikter voor het beschouwen van de zaak als geheel. Door de twee benaderingen te combineren zou het globale perspectief van scenario's in de context kunnen voorzien die nodig is om te bepalen welke variabelen relevant zijn voor een kansbenadering, terwijl de kansbenadering het benodigde hulpmiddel levert om de bewijskracht van elk bewijsstuk voor een scenario te analyseren.

Als kansbenadering zijn Bayesiaanse netwerken mogelijk een geschikt hulpmiddel voor het modelleren van een zaak als geheel, omdat ze goed kunnen omgaan met afhankelijkheden tussen hypotheses en variabelen in een complex model. Er is echter nog niet veel onderzoek gedaan naar het modelleren van gehele rechtszaken in een Bayesiaans netwerk, met uitzondering van de uitgebreide analyse van Kadane en Schum (1996). De constructie van een Bayesiaans netwerk voor een complexe zaak ligt niet voor de hand. Daarom is de eerste onderzoeksvraag van dit proefschrift:

Onderzoeksvraag 1. Hoe kunnen hulpmiddelen uit de scenariobenadering helpen bij de constructie van een Bayesiaans netwerk dat juridisch bewijs in een strafzaak modelleert?

Wanneer een Bayesiaans netwerk toegepast wordt op een gehele zaak is het bovendien cruciaal dat de rechter of jury begrijpt wat het netwerk precies modelleert. Met name moeten zij de aannames die ten grondslag liggen aan het model begrijpen, omdat dit is wat uiteindelijk de uitkomst bepaalt (Fenton and Neil, 2011) Daarom is de tweede onderzoeksvraag van dit proefschrift:

Onderzoeksvraag 2. Hoe kunnen hulpmiddelen uit de scenariobenadering helpen bij de uitleg van een Bayesiaans netwerk dat juridisch bewijs in een strafzaak modelleert?

Een methode voor de constructie en uitleg van Bayesiaanse netwerken

In dit proefschrift presenteren we een methode voor het construeren en uitleggen van Bayesiaanse netwerken met behulp van scenario's. Scenario's worden in een Bayesiaans netwerk gerepresenteerd, evenals het bewijs en de kwaliteit van de scenario's. Tegelijkertijd functioneren de scenario's als een richtlijn voor het construeren en uitleggen van het netwerk. Als uitleg van het netwerk worden de scenario's, scenariokwaliteit en bewijskracht gerapporteerd.

Constructie Deel II van dit proefschrift gaat over een methode voor de constructie van een Bayesiaans netwerk. Geïnspireerd op eerder werk van Hepler et al. (2004) en Fenton et al. (2013) gebruiken we netwerkfragmenten, genaamd idiomen, om scenario's in een Bayesiaans netwerk te modelleren. Zulke idiomen zijn te gebruiken als een soort kant en klaar sjabloon dat locaal de grafische structuur van het netwerk weergeeft. Vier idiomen voor het representeren van scenario's worden voorgesteld: het scenarioschema-idioom, het subscenario-idioom, het variatie-idiom en het samengevoegde-scenario's-idioom.

Het scenarioschema-idioom wordt gebruikt om de coherentie van een scenario te vangen, door het scenario als een cluster van elementen weer te geven, bijeen gehouden door een scenario-knoop, dat het scenario als geheel representeert. Bijbehorende kansen zijn deels vastgelegd door het scenarioschema-idioom en garanderen de overdracht van bewijskracht, waarbij een element van een scenario waarschijnlijker wordt doordat een ander element van dat scenario ondersteund wordt met bewijs, als resultaat van coherentie. Op vergelijkbare wijze kan het subscenario-idioom gebruikt worden om de coherentie van een subscenario te representeren. Met het variatie-idioom kunnen variaties binnen een scenario worden weergegeven en met het samengevoegde-scenario's-idioom worden verschillende scenario-structuren samengevoegd in één netwerk.

Uiteenlopende scenarioschema-idiomen kunnen worden gebruikt als kant en klare structuren om scenario's te representeren. Elk scenarioschema-idioom geeft de Bayesiaansenetwerkstructuur voor hoe een typisch scenario verloopt, zoals een inbraakschema voor een typische inbraak en een moordschema voor een typische moord. De scenarioschema-idiomen zijn geannoteerd, wat later wordt gebruikt voor de uitleg van het Bayesiaanse netwerk. De scenarioschema-idiomen versimpelen bovendien het constructieproces aangezien de onderliggende structuur van een scenario beschikbaar is als idioom.

Het scenarioschema-idioom maakt het bovendien mogelijk om de kwaliteit van een scenario te representeren, in termen van volledigheid, consistentie en plausibiliteit zoals voorgesteld door Pennington en Hastie (1993). In de voorgestelde methode is een scenario volledig wanneer het een scenarioschema-idioom past en completeert. Een scenario is consistent als het geen elementen bevat die samen inconsistent zijn en de plausibiliteit van een scenario kan worden afgelezen van de a priori kansverdeling van de scenario-knoop. Op vergelijkbare wijze is de plausibiliteit van een element in het scenario gegeven door de a priori kansverdeling voor dat element.

Met behulp van de vier idiomen en de interpretatie van scenariokwaliteit kan een Bayesiaans netwerk stapsgewijs geconstrueerd worden door een scenario uit te vouwen tot het gewenste detailniveau. Het concept van het uitvouwen van scenario's is gebaseerd op het idee dat elementen van een scenario altijd kunnen worden uitgebreid tot het gewenste niveau door ze te vervangen door een subscenario. Een constructiemethode in vijf stappen maakt het mogelijk om stapsgewijs een Bayesiaans netwerk te construeren dat verschillende alternatieve scenario's representeert. In Hoofdstuk 4 is de voorgestelde constructiemethode getest aan de hand van een echte casus.

Uitleg Hoewel de constructiemethode tamelijk grote en complexe netwerken kan opleveren, is de graaf geannoteerd en gestructureerd aan de hand van de vier idiomen. In het bijzonder zorgen de scenarioschema-idiomen ervoor dat de structuur van het netwerk dicht bij de oorspronkelijke scenario's blijft, doordat de knopen in het netwerk corresponderen met proposities in het scenario, en pijlen met verbanden. In Deel III van dit proefschrift wordt een techniek voorgesteld voor het extraheren van scenario's uit een Bayesiaans netwerk, met als resultaat een tekst voor elke scenariostructuur in het netwerk.

In aanvulling op de scenario's in tekst-vorm wordt bovendien een rapportage-format voorgesteld. Zo'n rapportage-format rapporteert welke scenario's in het netwerk gerepresenteerd zijn, evenals informatie over scenariokwaliteit en bewijskracht. We geven een maat voor bewijskracht van een bewijsstuk voor een scenario en observeren daarnaast dat scenariokwaliteit en in het bijzonder plausibiliteit beïnvloeden hoe de kans op een scenario verandert als gevolg van bewijskracht. Wanneer een element van een scenario in eerste instantie heel implausibel is, dan heeft bewijs voor dat element een sterker effect op de kans op het scenario als geheel dan bewijs voor een ander, plausibeler element in dat scenario. Om te begrijpen waarom een scenario een bepaalde kans heeft, is het op dezelfde wijze ook van belang om te kijken naar de implausibele elementen waar geen bewijs voor is. Deze verklaren immers waar nog cruciale bewijsstukken ontbreken. In het voorgestelde rapportageformat stellen we voor om implausibele elementen te rapporteren, zowel die elementen waar bewijs voor is als die elementen waar geen bewijs voor is.

De voorgestelde rapportage kan door een rechter of jury gebruikt worden om een Bayesiaans netwerk te begrijpen in termen van scenario's. Het rapport bevat alle ingrediënten die nodig zijn voor een scenariobenadering, maar dan voorzien van kansinformatie. Elk scenario wordt gerapporteerd met een a posteriorikans, de kwaliteit van elk scenario wordt gerapporteerd en de kracht van elke bewijsstuk voor elk scenario wordt kwalitatief gerapporteerd, evanals de gecombineerde kracht van al het bewijs. De technieken voor het uitleggen van Bayesiaanse netwerken zijn getest aan de hand van een casus in Hoofdstuk 6.

Onderzoeksresultaten

Het resultaat van het onderzoek in dit proefschrift is een methode voor het construeren en uitleggen van een Bayesiaans netwerk, zoals beschreven in de vorige sectie. In deze sectie bespreken we resultaten wat betreft de twee onderzoeksvragen.

Onderzoeksvraag 1: constructie Als onderdeel van de constructiemethode worden idiomen gebruikt om scenario's te representeren in een Bayesiaans netwerk. Doordat scenario's een beschrijving geven van wat er in een Bayesiaans netwerk gerepresenteerd moet worden, helpt het weergeven van verschillende alternatieve scenario's bij de constructie van een Bayesiaans netwerk. Bovendien voorziet het concept van een scenarioschema-idioom in kant en klare structuren voor typische misdaden, zoals inbraken of moorden. Het concept van een scenarioschema uit de scenariobenadering wordt op deze wijze gebruikt om de de constructie van een Bayesiaans netwerk te structureren.

In de Anjum-casus in Hoofdstuk 4 konden we constateren dat scenarioschema-idiomen inderdaad het modelleerproces structureren. Zelfs in het geval dat deze scenarioschema-idiomen tijdens het modelleren bedacht worden, helpt dit om de constructie minder ad hoc te maken dan zonder het gebruik van scenarioschema's, aangezien de modelleerder moet nadenken over de onderliggende structuur van een scenario.

Een scenarioschema-idioom is ook nodig om te zorgen dat scenariokwaliteit in een Bayesiaans netwerk kan worden weergegeven. Hoewel de representatie van scenariokwaliteit niet direct bijdraagt aan de vereenvoudiging van het constructieproces, kan nadenken over volledigheid, consistentie en in het bijzonder plausibiliteit toch helpen bij het vaststellen van kansen. In de Anjum-casus vonden we dat de kwaliteit van een scenario goed kan worden gerepresenteerd, maar dat het nog niet gemakkelijk is de plausibiliteit van een scenario en elementen in het scenario vast te stellen.

Ook de voorgestelde constructieprocedure maakt gebruik van concepten uit de scenariobenadering. De procedure maakt gebruik van het globale perspectief van scenario's en van het feit dat (sub)scenario's kunnen worden uitgevouwen tot het gewenste detailniveau. In de Anjum-casus bleek dat de constructieprocedure en vooral het uitvouwen van scenario's behulpzaam is voor het behouden van overzicht tijdens het werken aan een groot en complex netwerk. Echter, de casus liet ook zien dat de laatste stap in de constructieprocedure, waarin kansen moeten worden gespecificeerd, buitenproportioneel groot is.

Samenvattend zijn de bijdragen in dit proefschrift ten aanzien van de eerste onderzoeksvraag als volgt:

1. Vier idiomen voor het representeren van scenario's in een Bayesiaans netwerk;

2. Scenarioschema-idiomen als kant en klare structuren, gebaseerd op het concept van scenarioschema's;

3. een kans-interpretatie van scenariokwaliteit in volledigheid, consistentie en plausibiliteit;

4. een procedure voor de stapsgewijze constructie van een Bayesiaans netwerk waarin alternatieve scenario's, scenariokwaliteit en bewijs worden gemodelleerd;

5. een evaluatie van de methode aan de hand van een casus.

Onderzoeksvraag 2: uitleg Een netwerk dat gebouwd is met de constructiemethode uit dit proefschrift is geannoteerd en gestructureerd zodat later scenario's uit het netwerk geëxtraheerd kunnen worden. Op deze manier is het concept van scenarioschema's dat bij de constructie wordt gebruikt ook behulpzaam bij het uitleggen van het netwerk. Zoals in de Nijmegen-casus in Hoofdstuk 6 kon worden vastgesteld, zorgen de geëxtraheerde scenario's ervoor dat de inhoud van het model wordt uitgelegd. De alternatieve scenario's fungeren dan als alternatieve verklaringen van het bewijs. De casus toonde ook dat een uitleg in termen van scenario's de inhoud van het netwerk op hoog niveau goed overbrengt, maar weinig bevat over de locale structuur van het netwerk.

In aanvulling op de geëxtraheerde scenario's wordt ook scenariokwaliteit en bewijskracht gerapporteerd. Daardoor kan het rapport door een rechter of jury gebruikt worden om de zaak te analyseren met een scenariobenadering. Het rapport bevat alle ingrediënten voor zo'n scenariobenadering, aangevuld met kansinformatie. Deze rapportage is gebaseerd op ideeën uit de scenariobenadering, waaruit blijkt dat de scenario's, scenariokwaliteit en het bewijs samen voldoende zijn voor een analyse van de zaak. De Nijmegen-casus liet zien dat deze drie ingrediënten inderdaad inzicht geven in de zaak, aangezien de rapportage goed overeen kwam met de bevindingen in het hoger beroep van deze zaak.

Tenslotte observeerden we ook dat scenariokwaliteit en vooral plausibiliteit invloed hebben op hoe bewijskracht de kans op een scenario verandert. Plausibiliteit kan dus inzicht geven in waarom een scenario een hogere kans heeft dan een ander scenario. In het voorgestelde rapport worden implausibele elementen in een scenario dus gerapporteerd. Het concept van scenariokwaliteit uit de scenariobenadering wordt zo gebruikt bij het uitleggen van een Bayesiaans netwerk.

Samenvattend zijn de bijdragen van dit proefschrift ten aanzien van de tweede onderzoeksvraag als volgt:

1. Een procedure voor de extractie van scenario's in text-vorm uit een Bayesiaans netwerk dat met scenarioschema's is geconstrueerd;

2. een rapportageformat voor het rapporteren van scenario concepten op basis van een Bayesiaans netwerk in termen van scenario's, scenariokwaliteit en bewijskracht, voorzien van kansinformatie;

3. evaluatie van de voorgestelde uitlegmethode met een casus.

Conclusie

In dit proefschrift hebben we een methode gepresenteerd die scenario's als richtlijn gebruikt voor de constructie en uitleg van een Bayesiaans netwerk. Tegelijkertijd worden de alternatieve scenario's, hun kwaliteit en het bewijs samen in een Bayesiaans netwerk gemodelleerd. Bij wijze van uitleg worden deze scenario's, kwaliteit en bewijskracht gerapporteerd, vergezeld van kansinformatie.

Voor deze methode hebben we gebruik gemaakt van Bayesiaanse netwerken. Hoewel de rekenmethode in een Bayesiaans netwerk niet ter discussie staat, is er nog veel discussie mogelijk over hoe een situatie in de echte wereld, bijvoorbeeld een rechtszaak, correct gemodelleerd moet worden. Dit onderzoek was niet zozeer gericht op de kansen in het netwerk, maar vooral op de graafstructuur. Toch liepen we tegen het probleem aan dat veel getallen expliciet gemaakt moeten worden. Om die reden is er meer onderzoek nodig om te vinden hoe Bayesiaanse netwerken verantwoord kunnen worden gebruikt in de rechtszaal, voordat welke methode dan ook (inclusief die in dit proefschrift) klaar is om echt toegepast te worden.

Acknowledgements

Some important factors in the successful completion of this PhD-project were an interesting topic, a clear research proposal to work with and enthusiastic and supportive supervisors. Many thanks to Henry and Bart providing these excellent conditions to work with and for meeting me regularly to discuss my work, always being interested in my progress and providing me with valuable feedback. Also many thanks to Silja, who managed to always give such clear and concrete feedback despite the physical distance. And many thanks to Rineke, who seems to be a professional at reading and commenting on a text, even when the subject is not in her specific field of expertise.

During these four years, we have established a small project group, lovingly called the 'Groningen School of Legal Evidence'. I would like to thank Sjoerd and Floris, who worked on similar topics and were available for informal talks about our research every now and then. It has been very valuable to discuss things with someone who is not also your supervisor. I both wish you good luck with any future endeavours!

A PhD student working on one topic for four years cannot do without breaks. Thanks to all my fellow PhD students at ALICE for providing all kinds of breaks: lunch breaks, coffee breaks, afternoon walks and even the occasional board game break (I think that was only once?). A special thanks to Harmen for sharing an office with me during these four years, to Trudy for many conversations about our shared hobbies (or obsessions?) and to Jean-Paul for weekly squash games and just generally being a very good friend.

Many healthy distractions were also provided by my friends and family, for which I am very grateful. A particular shout out goes to Grupo Flamenco Rosario, for getting me out of my head and onto my feet!

Finally, I would like to thank some special people who have been of great support by just being in this world. My parents, Rutger, Anna and Ben: thank you so much for everything!

Bibliography

Aitken, C. (2012). Papers on the R v T debate [special issue]. *Law, Probability and Risk*, 11(4).

Anderson, T., Schum, D., and Twining, W. (2005). *Analysis of Evidence*. Cambridge University Press, Cambridge.

Anderson, T. and Twining, W. (1991). *The Analysis of Evidence: How to Do Things with Facts Based upon Wigmore's Science of Judicial Proof*. Aspen Publishers.

Association of Forensic Science Providers (2009). Standards for the formulation of evaluative forensic science expert opinion. *Science & Justice*, 49(3):161 – 164.

Barwise, J. and Etchemendy, J. (1999). *Language, proof and logic*. CSLI Publications.

Bench-Capon, T. and Dunne, P. (2007). Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641.

Bennett, W. and Feldman, M. (1981). *Reconstructing Reality in the Courtroom - Justice and Judgment in American Culture*. Methuen-Tavistock, London.

Berger, C. and Aben, D. (2010a). Bewijs en overtuiging: Een helder zicht op valkuilen (proof and conviction: A clear view of pitfalls). *Expertise en Recht*, 5/6:159 – 165.

Berger, C. and Aben, D. (2010b). Bewijs en overtuiging: Rationeel redeneren sinds Aristoteles (proof and conviction: Rational reasoning since Aristotle). *Expertise en Recht*, 2:52 – 56.

Berger, C. and Aben, D. (2010c). Bewijs en overtuiging: Redeneren in de rechtszaal (proof and conviction: Reasoning in the courtroom). *Expertise en Recht*, 3:86 – 90.

Bex, F. (2011). *Arguments, Stories and Criminal Evidence, a Formal Hybrid Theory*. Springer, Dordrecht.

Bex, F., Prakken, H., Reed, C., and Walton, D. (2003). Towards a formal account of reasoning about evidence: Argumentation schemes and generalisations. *Artificial Intelligence and Law*, 11:125 – 165.

Bex, F., van den Braak, S., van Oostendorp, H., Prakken, H., Verheij, B., and Vreeswijk, G. (2007). Sense-making software for crime investigation: How to combine stories and arguments? *Law, Probability and Risk*, 6(1-4):145–168.

Bex, F., van Koppen, P., Prakken, H., and Verheij, B. (2010). A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law*, 18:123–152.

Biedermann, A., Taroni, F., Delemont, O., Semadeni, C., and Davison, A. (2005). The evaluation of evidence in the forensic investigation of fire incidents (part i): An approach using Bayesian networks. *Forensic Science International*, 147(1):49–57.

Carnap, R. (1962). *Logical Foundations of Probability*. University of Chicago Press, Chicago, second edition.

Crombag, H., Horselenberg, R., van Koppen, P., and Zeles, G. (2009). *Twee mysterieuze schietpartijen (Two mysterious shootings)*. Boom Juridische uitgevers, Den Haag.

Crombag, H. and Israëls, H. (2008). *Moord in Anjum (Murder in Anjum)*. Boom Juridische uitgevers, Den Haag.

Crombag, H., van Koppen, P., and Wagenaar, W. (1992). *Dubieuze Zaken: De Psychologie van Strafrechtelijk Bewijs*. Uitgeverij Contact, Amsterdam.

Dawid, A. (2009). Beware of the DAG. In *Journal of Machine Learning Research: Workshop and Conference Proceedings*, volume 6, pages 59–86.

Dawid, A. and Evett, I. (1997). Using a graphical method to assist the evaluation of complicated patterns of evidence. *Journal of Forensic Sciences*, 42(2).

Dawid, A. P., Twining, W., and Vasiliki, M., editors (2011). *Evidence, Inference and Enquiry*. Oxford University Press, Oxford.

de Zoete, J., Sjerps, M., Lagnado, D., and Fenton, N. (2015). Modelling crime linkage with Bayesian networks. *Science & Justice*, 55(3):209–217.

Druzdzel, M. (1996). Qualitative verbal explanations in Bayesian belief networks. *Artificial Intelligence and Simulation of Behaviour Quarterly*, 94:43–54.

Evett, I. and Weir, B. (1998). *Interpreting DNA evidence: Statistical genetics for forensic scientists*. Sinauer.

Fenton, N. (2011). Science and law: Improve statistics in court. *Nature*, 479:36–37.

Fenton, N., Berger, D., Lagnado, D., Neil, M., and Hsu, A. (2014). When 'neutral' evidence still has probative value (with implications from the Barry George Case). *Science & Justice*, 54(4):274 – 287.

Fenton, N. and Neil, M. (2011). Avoiding probabilistic reasoning fallacies in legal practice using Bayesian networks. *Australian Journal of Legal Philosophy*, 36:114 – 151.

Fenton, N., Neil, M., and Lagnado, D. (2011). Modelling mutually exclusive causes in Bayesian networks. Available online: www.eecs.qmul.ac.uk/∼norman/papers/soft_evidencence_mutual_exclusivity.pdf.

Fenton, N., Neil, M., and Lagnado, D. (2013). A general structure for legal arguments using Bayesian networks. *Cognitive Science*, 37:61 – 102.

Henrion, M. and Druzdzel, M. (1990). Qualitative propagation and scenario-based approaches to explanation of probabilistic reasoning. *Uncertainty in Artificial Intelligence*, 6:17–32.

Hepler, A., Dawid, A., and Leucari, V. (2004). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability & Risk*, 6:275–293.

Israëls, H. (2006). *Moord in Wamel (Murder in Wamel)*. Boom Juridische uitgevers, Den Haag.

Jensen, F. and Nielsen, T. (2007). *Bayesian Networks and Decision Graphs*. Springer, New York.

Kadane, J. and Schum, D. (1996). *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. John Wiley and Sons, New York.

Kaptein, H., Prakken, H., and Verheij, B. (2009). *Legal Evidence and Proof: Statistics, Stories, Logic*. Ashgate Publishing Company, Aldershot.

Keppens, J. (2011). On extracting arguments from Bayesian network representations of evidential reasoning. In Ashley, K. and van Engers, T., editors, *The 13th International Conference on Artificial Intelligence and Law*, pages 141 – 150, New York. ACM.

Keppens, J. (2012). Argument diagram extraction from evidential Bayesian networks. *Artificial Intelligence and Law*, 20(2):109 – 143.

Keppens, J. and Schafer, B. (2006). Knowledge based crime scenario modelling. *Expert Systems with Applications*, 30(2):203–222.

Kjærulff, U. and Madsen, A. (2008). *Bayesian Networks and Influence Diagrams*. Springer, New York.

Koiter, J. (2006). Visualizing inference in Bayesian networks. Master's thesis, Delft University of Technology.

Koller, D. and Pfeffer, A. (1997). Object-oriented Bayesian networks. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, pages 302–313, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Lacave, C. and Díez, F. (2002). A review of explanation methods for Bayesian networks. *The Knowledge Engineering Review*, 17(2):107–127.

Lacave, C., Onisko, A., and Díez, F. (2001). Debugging medical Bayesian networks with Elviras explanation capability. In *Proceedings of the Workshop on Bayesian Models in Medicine, Eighth European Conference on Artificial Intelligence in Medicine*, pages 47–52.

Lagnado, D., Fenton, N., and Neil, M. (2012). Legal idioms: A framework for evidential reasoning. *Argument and Computation*, 4(1):46–63.

Laskey, K. and Mahoney, S. (1997). Network fragments: Representing knowledge for constructing probabilistic models. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, pages 334–341. Morgan Kaufmann Publishers Inc.

MacCrimmon, M. and Tillers, P. (2002). *The Dynamics of Judicial Proof*. Physica-Verlag, Heidelberg.

Nordgaard, A., Ansell, R., Drotz, W., and Jaeger, L. (2012). Scale of conclusions for the value of evidence. *Law, Probability and Risk*, 11(1):1–24.

Pardo, M. and Allen, R. (2008). Juridicial proof and the best explanation. *Law and Philosophy*, 27(3):223–268.

Pennington, N. and Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, 51(2):242 – 258.

Pennington, N. and Hastie, R. (1988). Explanation-based decision making: Affects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14:521–533.

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. (1993). *Inside the Juror*, chapter The Story Model for juror decision making, pages 192 – 221. Cambridge University Press, Cambridge.

Renooij, S. (2001). Probability elicitation for belief networks: Issues to consider. *The Knowledge Engineering Review*, 16(3):255 – 269.

Rumelhart, D. (1975). Notes on a schema for stories. In Bobrow, D. and Collins, A., editors, *Representation and Understanding: Studies in Cognitive Science*. Academic Press, New York, New York.

Schank, R. and Abelson, R. (1977). *Scripts, Plans, Goals and Understanding, An Inquiry into Human Knowledge Structures*. Lawrence Erlbaum, Hillsdale.

Scherrenburg, O., van Beek, M., Jonker, S., Goldstoff, N., Hardeman, M., Luten, W., Lagerweij, S., Luijten, A., Diesfeldt, A., Das, A., and van Koppen, P. (2013). *De Moddermoord (The mud murder)*. Boom Juridische uitgevers, Den Haag.

Schum, D. A. (2001). Alternative views of argument construction from a mass of evidence. *Cardozo Law Review*, 22:1461–1502.

Shen, Q., Keppens, J., Aitken, C., Schafer, B., and Lee, M. (2006). A scenario-driven decision support system for serious crime investigation. *Law, Probability and Risk*, 5(2):87–117.

Sileno, G., Boer, A., and van Engers, T. (2014). Legal knowledge conveyed by narratives: Towards a representational model. In Finlayson, M., Meister, J., and Bruneau, E., editors, *2014 Workshop on Computational Models of Narrative*, volume 41, pages 182–191, Dagstuhl, Germany. Schloss Dagstuhl.

Suermondt, H. (1992). *Explanation in Bayesian Belief Networks*. PhD thesis, Stanford University.

Taroni, F., Aitken, C., Garbolino, P., and Biedermann, A. (2006). *Bayesian Networks and Probabilistic Inference in Forensic Science*. Wiley, Chichester.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, 12:435–502.

Thagard, P. (2004). Causal inference in legal decision making: Explanatory coherence versus Bayesian networks. *Applied Artificial Intelligence*, 18(3):231 – 249.

Thompson, W. C. and Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, 11:167–187.

Timmer, S., Meyer, J., Prakken, H., Renooij, S., and Verheij, B. (2013). Inference and attack in Bayesian networks. In Hindriks, K., de Weerdt, M., van Riemsdijk, B., and Warnier, M., editors, *Proceedings of the 25th Benelux Conference on Artificial Intelligence*, pages 199 – 206.

Timmer, S., Meyer, J., Prakken, H., Renooij, S., and Verheij, B. (2014). Extracting legal arguments from forensic Bayesian networks. In Hoekstra, R., editor, *The 27th International Conference on Legal Knowledge and Information Systems (JURIX) 2014*, pages 71–80, Amsterdam. IOS Press.

Timmer, S., Meyer, J., Prakken, H., Renooij, S., and Verheij, B. (2015a). A structure-guided approach to capturing Bayesian reasoning about legal evidence in argumentation. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law*, pages 109– 118, New York. ACM Press.

Timmer, S., Meyer, J.-J. C., Prakken, H., Renooij, S., and Verheij, B. (2015b). Capturing critical questions in Bayesian network fragments. In *Legal Knowledge and Information Systems. JURIX 2015: The Twenty-eighth Annual Conference*, pages 173–176.

Timmer, S., Meyer, J.-J. C., Prakken, H., Renooij, S., and Verheij, B. (2015c). Explaining Bayesian networks using argumentation. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 13th European Conference, ECSQARU 2013, Compiègne, France, July 15-17, 2015. Proceedings*, pages 83–92. Springer.

Timmer, S., Meyer, J.-J. C., Prakken, H., Renooij, S., and Verheij, B. (2015d). Explaining legal Bayesian networks using support graphs. In *Legal Knowledge and Information Systems. JURIX 2015: The Twenty-eighth Annual Conference*, pages 121–130.

van den Braak, S. (2010). *Sensemaking software for crime analysis*. PhD thesis, Utrecht University.

van den Braak, S., Vreeswijk, G., and Prakken, H. (2007). AVERs: An argument visualization tool for representing stories about evidence. In *Proceedings of the 11th international conference on Artificial intelligence and law*, pages 11–15. ACM.

van Eemeren, F., Garssen, B., Krabbe, E., Henkemans, A. S., Verheij, B., and Wagemans, J. (2014). *Handbook of Argumentation Theory*. Springer.

van Gosliga, S. and van de Voorde, I. (2008). Hypothesis management framework: A flexible design pattern for belief networks in decision support systems. In *6th Bayesian Modelling Applications Workshop on UAI 2008, Helsinki, Finland*.

van Koppen, P. (2011). *Overtuigend Bewijs (Convincing Evidence)*. Nieuw Amsterdam, Amsterdam.

Verheij, B. (2012). Integrating argumentation, narrative and probability in legal evidence. In Finlayson, M., editor, *The Third Workshop on Computational Models of Narrative*, pages 176–177.

Verheij, B. (2014). To catch a thief with and without numbers: Arguments, scenarios and probabilities. *Law, Probability & Risk*, 13(3-4):307–325.

Verheij, B., Bex, F., Timmer, S., Vlek, C., Meyer, J.-J., Renooij, S., and Prakken, H. (2016). Arguments, scenarios and probabilities: Connections between three normative frameworks for evidential reasoning. *Law, Probability & Risk*, 15(1):35–70.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013a). Modeling crime scenarios in a Bayesian network. In *Proceedings of the 14th International Conference on Artificial Intelligence and Law*, pages 150 – 159, New York. ACM Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013b). Representing and evaluating legal narratives with subscenarios in a Bayesian network. In Finlayson, M., Fisseni, B., Löwe, B., and Meister, J., editors, *2013 Workshop on Computational Models of Narrative*, pages 315–332, Saarbrücken/Wadern, Germany. Schloss Dagstuhl.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013c). Unfolding crime scenarios with variations: A method for building Bayesian networks for legal narratives. In Ashley, K., editor, *Legal Knowledge and Information Systems: JURIX 2013: The Twenty-Sixth Annual Conference*, pages 145 – 154, Amsterdam. IOS Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014a). Building Bayesian networks for legal evidence with narratives: A case study evaluation. *Artificial Intelligence and Law*, 22(4):375 – 421.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014b). Extracting scenarios from a Bayesian network as explanations for legal evidence. In Hoekstra, R., editor, *The 27th International Conference on Legal Knowledge and Information Systems (JURIX) 2014*, pages 103–112, Amsterdam. IOS Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2015a). Constructing and understanding Bayesian networks for legal evidence with scenario schemes. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law*, pages 128– 137, New York. ACM Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2015b). Representing the quality of crime scenarios in a Bayesian network. In Rotolo, A., editor, *Legal Knowledge and Information Systems. JURIX 2015: The Twenty-eighth Annual Conference*, pages 131 – 140, Amsterdam. IOS Press.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law*. In press.

Wagenaar, W., van Koppen, P., and Crombag, H. (1993). *Anchored Narratives: The Psychology of Criminal Evidence.* Harvester Wheatsheaf, Hemel Hempstead.

Wellman, P. and Henrion, M. (1993). Explaining 'explaining away'. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):287–292.

Wigmore, J. (1931). *The Principles of Judicial Proof or the Process of Proof as Given by Logic, Psychology, and General Experience, and Illustrated in Judicial Trials.* Little, Brown and Company, Boston.

Index