

# Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning

Bart Verheij<sup>a</sup>   Floris Bex<sup>b</sup>   Sjoerd T. Timmer<sup>b</sup>   Charlotte S. Vlek<sup>a</sup>  
John-Jules Ch. Meyer<sup>b</sup>   Silja Renooij<sup>b</sup>  
Henry Prakken<sup>b,c</sup>

<sup>a</sup>*Institute of Artificial Intelligence, University of Groningen*

<sup>b</sup>*Department of Information and Computing Sciences, Utrecht University*

<sup>c</sup>*Faculty of Law, University of Groningen*

## Abstract

Due to the uses of DNA profiling in criminal investigation and decision-making, it is ever more common that probabilistic information is discussed in courts. The people involved have varied backgrounds, as fact-finders and lawyers are more trained in the use of non-probabilistic information, while forensic experts handle probabilistic information on a routine basis. Hence, it is important to have a good understanding of the sort of reasoning that happens in criminal cases, both probabilistic and non-probabilistic. In the present paper, we report results on combining three normative reasoning frameworks from the literature: arguments, scenarios and probabilities. We discuss a hybrid model that connects arguments and scenarios, a method to probabilistically model possible scenarios in a Bayesian network, a method to extract arguments from a Bayesian network, and a proposal to model arguments for and against different scenarios in standard probability theory. These results have been produced as parts of research projects on the formal and computational modelling of evidence. The present paper reviews these results, shows how they are connected and where they differ, and discusses strengths and limitations.

## 1 Introduction

DNA profiling has become a standard tool in criminal investigation and decision making. A good DNA profile match has a high information value, and the associated statistics have a solid and well-understood scientific foundation. As a result of the success of DNA evidence, the interpretation of statistical information—whether presented numerically or not—has become a common task for fact-finders and decision-makers in criminal cases, such as judges and juries. These fact-finders are typically used to non-numeric styles of reasoning involving arguments and scenarios, whereas the forensic experts presenting DNA evidence in courts are trained in numeric reasoning in terms of probabilities and statistics. Hence, miscommunication and misinterpretation is a real danger. Furthermore, cognitive scientists have shown experimentally that people are prone to many kinds of reasoning errors, with errors in probabilistic reasoning among the most notorious (Kahneman, 2011; Thompson and Schumann, 1987; Thompson, 2013). These findings are also relevant outside the experimental lab, now that probabilistic reasoning is associated with a number of infamous miscarriages of justice, among them the Lucia de Berk

and Sally Clark cases.<sup>1</sup> Interestingly, in both the Lucia de Berk and Sally Clark cases, it were the statistically trained expert witnesses who played a pivotal role in the erroneous probabilistic reasoning that led to the wrong decisions.

The prevention of reasoning errors requires a generally accepted and generally applicable normative framework that can be used to establish whether reasoning is correct or not. In the area of criminal fact-finding, a generally accepted and generally applicable normative framework does not exist. In the literature, three types of normative framework for reasoning from evidence to facts are available. The types can be distinguished by their emphasis on arguments, scenarios or probabilities (Kaptein et al., 2009; Dawid et al., 2011; Anderson et al., 2005).

*Argumentative approaches* to evidential reasoning emphasise the arguments based on evidence that support or attack conclusions. For instance, the conclusion that someone died of natural causes can be supported with the argument based on the coroner's report which states that the death was due to heart failure. The conclusion can be attacked when a suicide note is found. Argumentative approaches go back to John Henry Wigmore's work at the start of the 20th century (Wigmore, 1913). His diagrams of the structure of the arguments used in evidential reasoning are precursors of the diagrams in today's argument mapping software tools (Kirschner et al., 2003; Verheij, 2005). The theory of argumentation is actively studied, both formally and non-formally (van Eemeren et al., 2014b). In recent years, the computational study of argumentation has become a lively field of research (Bench-Capon and Dunne, 2007; Rahwan and Simari, 2009).

In *scenario approaches*, the evidence is interpreted in terms of scenarios that give a coherent account of what happened before, during and after the crime. For instance, when someone dies and a suicide note is found, the person's death may be explained by a scenario in which the person has killed himself by taking an overdose of sleeping pills, following a period of depression. An alternative scenario, e.g., that the person died of natural causes, can also be used to explain the death, but this alternative does not explain why a suicide note was found. Thus, scenario approaches are often a type of inference to the best explanation (Pardo and Allen, 2008). Whilst research has shown that scenarios are a natural way to make sense of the evidence in a complex case (Pennington and Hastie, 1993), experiments showed the risk of the use of scenarios in courts, as it turned out that a well-constructed, but fictional scenario was more easily believed than a true story, that was not constructed well (Bennett and Feldman, 1981). Others emphasised the normative value of scenario analyses, especially as a means for the prevention of neglecting relevant scenarios (tunnel vision) (Wagenaar et al., 1993). Recently, the relevance of scenarios in evidential reasoning has become the subject of computational research (Bex, 2011; Bex et al., 2010).

In *probabilistic approaches*, reasoning with evidence is studied using the mathematical language of the probability calculus. For instance, an expert can report that the odds of dying by suicide to dying by diseases of the circulatory system (such as heart failure) are about 1 to 20.<sup>2</sup> In recent years, probabilistic information has become prominent in courts by the evidential value of DNA profiling techniques. Since it can be hard for fact-finders (such as judges or juries) to interpret evidence of a probabilistic nature and to combine this evidence with other, often non-numeric, kinds of evidence, efforts have been aimed at developing verbal reporting techniques for probabilistic information suitable for the forensic domain (Evetts et al., 2000; Broeders, 2009).

---

<sup>1</sup>Both cases have Wikipedia pages: [http://en.wikipedia.org/wiki/Lucia\\_de\\_Berk](http://en.wikipedia.org/wiki/Lucia_de_Berk), [http://en.wikipedia.org/wiki/Sally\\_Clark](http://en.wikipedia.org/wiki/Sally_Clark). See also Buchanan (2007); Derksen and Meijding (2009); Schneps and Colmez (2013).

<sup>2</sup>E.g., the Dutch Central Bureau for Statistics reported that, in 2012, of a total of 843.5 deaths, 10.5 thousand people died of suicide (about 1 in 80) and 229.9 thousand of diseases of the circulatory system (about 1 in 4). (Source: 'Table: Health, lifestyle, health care use and supply, causes of death; from 1900'. Accessed online at [cbs.nl](http://cbs.nl); October 27, 2014.)

	Arguments	Scenarios	Probabilities
Adversarial setting	+	+/-	-
Global coherence	-	+	-
Degrees of uncertainty	+/-	-	+
Standard formalisation	+/-	-	+

Table 1: Characteristics of the three normative frameworks (Verheij, 2014b)

Each of the three normative perspectives has characteristic associated normative maxims. For instance, in argumentative approaches, it is necessary to consider all arguments for and against a position, since each additional argument can shift the balance in a case. In scenario approaches, it is necessary to consider all possible scenarios, lest we run the risk of so-called ‘tunnel-vision’: focusing on one or a few scenarios and thus neglecting other—possibly true—scenarios. In probabilistic approaches, it is necessary to follow the probability calculus, for instance by adhering to the formal connection between a conditional probability  $P(H|E)$  and its transposition  $P(E|H)$ , provided by Bayes’ theorem  $P(H|E)/P(H) = P(E|H)/P(E)$ .

The three normative perspectives emphasise different styles of analysing evidential reasoning. In argumentative approaches, the emphasis is on the adversarial setting of evidential reasoning, focusing on the pros and cons of the positions presented. In scenario approaches, the emphasis is on the globally coherent interpretation of the evidence available, in terms of explanatory scenarios. In probabilistic approaches, the emphasis is on the uncertainty of evidential reasoning, and how that uncertainty comes in degrees that can be measured numerically.

There are also differences in the formal development of the three normative perspectives. Probabilistic approaches have a widely accepted formalisation in terms of the standard probability calculus (although variations have been defended, and there are debates about the interpretation of probabilities; Hájek, 2011). Scenario approaches are formally less well-developed, with the most elaborated proposals connecting scenario approaches to argumentative (Bex, 2011) or probabilistic approaches (Shen et al., 2006). In recent decades, the formal study of argumentative approaches has received considerable attention, especially following Dung’s influential formal paper in 1995, but no single commonly accepted normative framework exists, now that Dung provides several so-called argumentation semantics.

Table 1 provides an overview of characteristics of the three normative perspectives. The adversarial setting, global coherence and degrees of uncertainty are the characteristic modelling strengths of argumentative, scenario and probabilistic approaches, respectively. These strengths correspond to a + entry in the table. Some entries indicate a +/- . For instance, scenario approaches also point to the adversarial setting by their inclusion of alternative scenarios, and some argumentative approaches consider argument strength, which is connected to degrees of uncertainty.

Since each of the normative frameworks focuses on evidence and proof, it is natural to consider how the central concepts of one framework are connected to those of another. For instance, each scenario in a scenario analysis can be regarded as a position supported or attacked by the arguments in an argumentative analysis. Also a new piece of evidence can flip the probabilistic odds of two competing hypothetical scenarios. Such evidence can hence be regarded as providing an argument supporting one scenario, and attacking another.

In this paper, we report on research on such connections between the three normative frameworks. The research reported on is being performed in the NWO Forensic Science research project ‘Designing and Understanding Forensic Bayesian Networks with Arguments and Scenarios’, in connection with the results from a previous project (NWO ToKeN project ‘Making sense of evidence’). Figure 1 shows the three normative frameworks, and their possible connections.

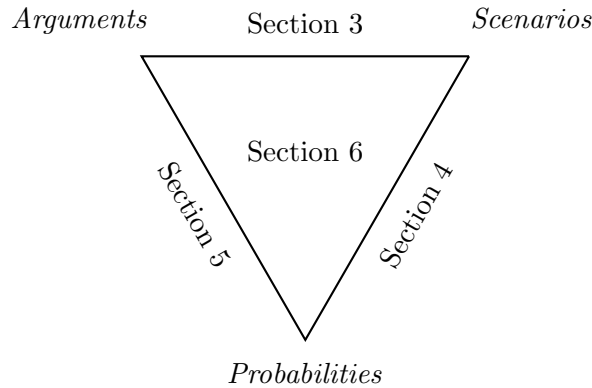


Figure 1: Three normative frameworks for evidential reasoning, with possible connections

We first discuss pairwise connections between arguments, scenarios and probabilities (Sections 3–5), and then suggest a view on arguments to and from scenarios in the context of probability theory (Section 6). In Section 3, connections between argumentative and scenario approaches are presented. This section highlights developments in the hybrid argumentative-narrative theory of evidential reasoning (Bex, 2014, 2011; Bex et al., 2010). Section 4 discusses connections between scenario and probabilistic approaches. The focus is on Bayesian networks, a kind of probabilistic graphical models that combine qualitative and quantitative information (Jensen and Nielsen, 2007). In the section, the embedding of scenarios in a Bayesian network is investigated, reporting on Vlek’s recent work (Vlek et al., 2013, 2014a,b). Section 5 addresses connections between argumentative and probabilistic approaches, focusing on recent work on extracting arguments from a Bayesian network (Timmer et al., 2013, 2014). In Section 6, connections between all three approaches to evidential reasoning are discussed, following work on the probabilistic modeling of arguments for and against scenarios (Verheij, 2014b,a). Sections 3 to 6 have different formal backgrounds. Sections 3 and 5 build on the argumentation formalism ASPIC+ (Prakken, 2010), Sections 4 and 5 on Bayesian networks, and Section 6 on standard probability theory and its underlying classical logic. Structurally, each of Sections 3 to 6 starts with a brief motivation and ends with the strengths and limitations of the findings presented. In Section 7, we conclude the paper with a summary. We now continue with a discussion of the running example used throughout the paper, and provide some background information about argumentation, scenarios and probabilities.

## 2 Background and running example

The research results discussed in this paper will be discussed using a running example, thereby allowing a good comparison of similarities and differences. We have selected a Dutch criminal case that was recently in the news.<sup>3</sup> The case is used for illustrative purposes only, and we do not intend to do justice to all elements of the decision.

In the case, a 16-year old girl was found dead and mutilated in a meadow. Physical evidence showed that Mary—not her real name—had been raped. She had been strangled, and her throat had been cut. Initially, there were no clear clues about who committed the crime. As there was an asylum seekers’ residence center in the area, some assumed that the brutal crime was committed by certain inhabitants of the center. Two asylum seekers, one from Iraq, one from Afghanistan, were considered as suspects, but were exonerated on the basis of DNA profiling.

<sup>3</sup>The district court’s decision is available at `rechtspraak.nl` with identifier NJFS 2013/155.

More than a decade after the crime, in part by continued media attention, an extensive screening of the local population was performed. A new law had just established such extensive screening as an investigative method in severe crimes. More than 8000 men were asked to provide a DNA sample. Apart from social pressure, there was no obligation to participate in the screening. The investigation's goal was to use Y-chromosome profiles to establish kinship relations between the perpetrator and the investigated men, thereby perhaps pointing to a limited set of persons to consider as possible suspects. Unexpectedly, a direct match was found, leading to the arrest of a then 45-year old suspect. The suspect had voluntarily participated in the screening. John—also a fictitious name—confessed to the crime, providing detailed knowledge of the crime's circumstances. John was convicted to 18 years imprisonment for rape and murder.

By the following properties, the selected case can be used for the comparison of the three normative frameworks for evidential reasoning:

1. **Probabilistic and non-probabilistic information.** *During the investigation, both probabilistic and non-probabilistic information is used.*

The probabilistic information used in the example case was based on DNA analysis. The probability that the DNA profile of a random male matches the main DNA profile of the blood trace found on the victim's coat (after his arrest) was estimated to be about 1 in 1500 billion billion (according to the court decision).<sup>4</sup>

Non-probabilistic information used in the case was for instance based on the suspect's confession that contained specific details about the crime's circumstances, and on physical traces such as the victim's body found in the meadow.

2. **Multiple scenarios.** *During the investigation of the case, different possible scenarios about what has happened are considered.*

In the selected case, three hypothetical scenarios can be distinguished: the crime was committed by an Iraqi asylum seeker, the crime was committed by an Afghani asylum seeker, the crime was committed by John.

3. **Arguments for and against events and scenarios.** *In the investigation, events and scenarios are both supported and attacked by arguments based on the evidence.*

In the selected case, arguments supporting the hypothesis that the suspect committed the crime were based on the direct DNA match found, and then corroborated by the suspect's confession that showed extensive specific perpetrator's knowledge.

Arguments against the scenarios that one of the two asylum seekers from Iraq and Afghanistan committed the crime were based on DNA evidence, that led to the exclusion of the two asylum seeker scenarios.

Using elements from the running example, we now illustrate the essence of the three normative perspectives on reasoning with evidence: arguments, scenarios and probabilities.

## 2.1 Arguments

The argumentative approach to evidential reasoning has its roots in Wigmore's evidence charts (Wigmore, 1913), which were later adapted for the modern age by Anderson et al. (2005) and rendered as formal defeasible arguments by Bex et al. (2003). An argumentative approach starts with the evidence (e.g., police reports, witness testimonies, forensic reports),

---

<sup>4</sup>We kept the original wording, repeating 'billion' ('De kans ... is ongeveer 1 op 1500 miljard miljard'). Below this number will appear as  $0.66 \cdot 10^{-21}$ .

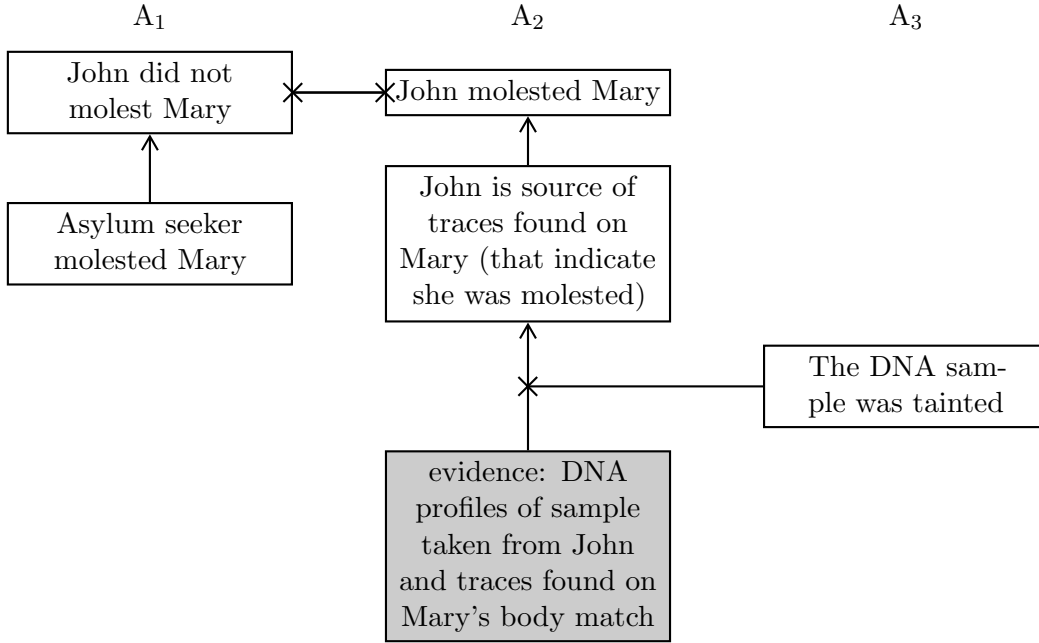


Figure 2: Evidential arguments that attack each other

and then consecutively applies evidential rules (e.g., ‘Witness testimony  $e$  is evidence for claim  $c$ ’), thus reasoning towards the ultimate claim in a case (e.g., ‘It was John who molested and killed Mary’). In the following, the example arguments in Figure 2 are used as an illustration.

Take, for example, the evidence ‘The DNA profiles of the sample taken from John and the traces found on Mary’s body match’, which can be used to support that ‘John is the source of the traces found on Mary’s body’ by applying the general rule that ‘If a DNA profile taken from person  $x$  matches trace  $y$ , this is evidence for the fact that  $x$  is probably the source of  $y$ ’. Such an argument is defeasible, and can for instance be attacked in case of a low prior probability (cf. the prosecutor’s fallacy; Thompson and Schumann, 1987).

Simple arguments can be chained and combined—for example, the conclusion of one argument can serve as a premise for another argument (see the middle part  $A_2$  in Figure 2, where the arrows denote inferences based on evidential rules)—and thus more complex arguments can be built to support claims in a case (here ‘John molested Mary’). As such, the argumentative approach focuses on how specific, single conclusions are based on the evidence.

The inferences in an argument are made using evidential inference rules of the form  $e$  is evidence for  $c$ , which act as a warrant for the inference (cf. Toulmin’s terminology, 1958). Such inference-warranting rules can range from very general—for example, ‘A person who is in a position to know about something can usually be believed’—, to more specific—for example, ‘A DNA expert who reports on a DNA matching can usually be believed’. Walton et al. (2008) have collected a large set of such rules that can act as inference warrants, referring to them as *argumentation schemes*, a term that has taken root in the field of formal and computational modelling of argument (Verheij, 2009).

Argumentation schemes can be used to determine whether an argument has all its necessary elements. Take, for example, the common scheme from expert opinion, which says that if person  $p$  says  $x$  and  $p$  is an expert on  $x$ , then we can believe  $x$ . If we now have an argument based on expert opinion which does not indicate that  $p$  was an expert—e.g., ‘Person  $p$  says  $x$ . Therefore we can believe  $x$ .’— we can say that the argument is incomplete, and is an enthymematic version of an extended argument—‘Person  $p$  says  $x$ . Person  $p$  is an expert on  $x$ . Therefore we can

believe  $x$ .’

In the argumentation literature, it is understood that arguments based on argumentation schemes are typically subject to exceptions, and do not guarantee the conclusions based on them under all circumstances. In the jargon: arguments are defeasible. It is therefore customary to define for each argumentation scheme some typical sources of doubt. These doubts, when phrased as critical questions, can then be used in the adversarial process of reasoning with evidence to probe and test the arguments. For example, a critical question for the expert opinion scheme is: ‘What do other experts in the field say?’ This question was asked and answered multiple times in the Lucia de Berk case, as the initial statistical analysis was repeatedly questioned and attacked by other experts.

The critical questions point to an important feature of argumentation, namely that it is adversarial or dialectical: not only arguments for a particular conclusion have to be considered, but also the relevant counterarguments. These counterarguments can have an opposite conclusion—we say that two such arguments rebut each other. Consider the example in Figure 2. From the premise that an asylum seeker molested Mary we can infer that it was not John who molested Mary (assuming that Mary was not molested by the asylum seeker and John together). Thus,  $A_1$  and  $A_2$  in Figure 2 rebut each other, indicated by the cross-headed arrow between ‘John molested Mary’ and ‘John did not molest Mary’.

Arguments can also be countered by an argument that gives an exception to the evidential rule that was used. For example, we normally accept the rule that a DNA match allows us to say something about the source of certain traces, allowing us to infer ‘John is the source of the traces found on Mary’s body’ from the evidence ‘The DNA profiles of the sample taken from John and the traces found on Mary’s body match’ ( $A_1$  in Figure 2). However, if the DNA sample was tainted by being handled improperly, we can argue that in this case we have an exception to the general reliability of DNA profiling, and that we cannot reasonably say something about the source of the trace on the basis of this evidence. In such a case, we say that one argument undercuts another: it is not a premise or conclusion that is denied, but rather the inference from a premise to a conclusion. In Figure 2, this is visualised by argument  $A_3$  attacking the first inference step of  $A_2$ .

When arguments can attack each other, it is not always clear which conclusions follow from them. The arguments in Figure 2 can be used as an illustration. If we only consider  $A_1$  and  $A_2$ , we have a conflict of arguments that is not resolved: there is support for the claim that John did not molest Mary ( $A_1$ ) and also for the claim that he did ( $A_2$ ). No clear conclusion follows about the issue whether John molested Mary or not. If it now turns out that the DNA sample was tainted ( $A_3$ ), the conflict is resolved, since  $A_2$  does no longer support that John molested Mary. Given the arguments in Figure 2, there is only support for the claim that John did not molest Mary, based on the premise that an asylum seeker did.

The evaluation of arguments that combine support and attack has been extensively studied in the literature. The arguments in Figure 2 can be handled by formalisms such as (Prakken, 2010; Bondarenko et al., 1997; Verheij, 2003b) and related argument diagramming software (Gordon et al., 2007; Verheij, 2003a). These formal and computational models build on the influential work by Pollock on defeasible argumentation and Dung on argument attack (Pollock, 1987, 1995; Dung, 1995) (see van Eemeren et al. (2014a) for an overview).

## 2.2 Scenarios

The scenario approach to evidential reasoning, also called the story-based or narrative approach, stems from legal psychology (Bennett and Feldman, 1981; Pennington and Hastie, 1993; Wagenaar et al., 1993). It has only relatively recently been further specified in both a formal setting (Bex,

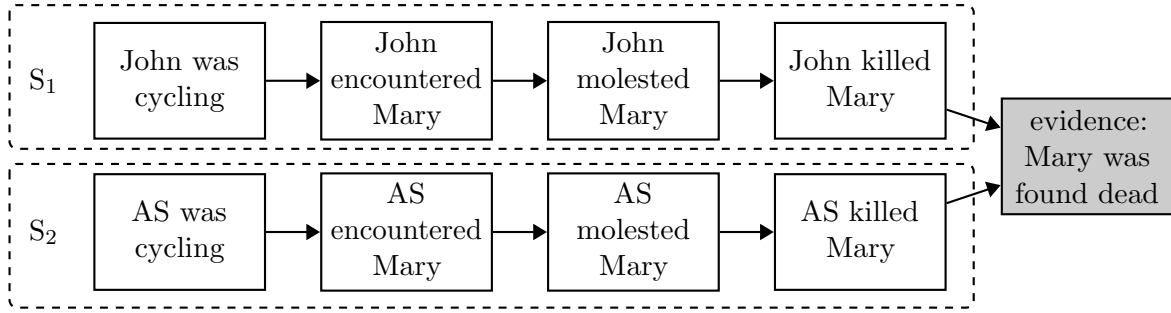


Figure 3: Two scenarios from the example case explaining evidence

2011; Bex et al., 2010) and a legal setting (Pardo and Allen, 2008). This approach focuses on scenarios or stories about what happened in a case (e.g., ‘John went cycling, encountered Mary and then molested and killed her’). These hypothetical scenarios—coherent sequences of events connected by (sometimes implicit) causal links of the form *c is a cause for e*—are then used to explain the evidence.

Take, for example, scenario  $S_1$  in Figure 3 (the arrows indicate the causal links between events), which includes the event ‘John killed Mary’. Now, with the causal rule ‘person  $x$  killing person  $y$  will cause  $y$  to die’ (the final link in the sequence), we can explain the evidence ‘Mary was found dead’. Note how the scenario approach focuses on the scenarios as-a-whole (‘John went cycling ... he ran into Mary ... he molested Mary ... he killed Mary’) rather than on a specific conclusion or main claim (‘It was John who molested and killed Mary’), which is the focus of an argumentative approach.

Like the argumentative approach, the scenario approach has an adversarial element: alternative or contradictory scenarios have to be compared. For example, the evidence ‘Mary was found dead’ is also explained by an alternative scenario  $S_2$  that an Iraqi man from the nearby asylum seekers’ residence center killed Mary. Assuming that John and the Iraqi man did not kill Mary together, the alternative scenario contradicts the main scenario in the case. An alibi scenario (e.g., that John was at home at the time of the murder), does not explain the evidence ‘Mary was found dead’ but it does counter the scenario  $S_1$  from Figure 3: John could not have been both at home and at the scene of the crime at the same time.

As mentioned before, the scenario approach is a kind of inference to the best explanation (Pardo and Allen, 2008; Josephson and Josephson, 1996): given the evidence, coherent scenarios should be constructed that *explain* the evidence. In addition to the explanatory power of stories, it is also possible to use stories to *predict* the (possible) existence of certain evidence. For example, in our example case, scenario  $S_1$  contains the event ‘John molested Mary’. If this were to be true, then it can be expected that biological traces (hair, sperm) of John can be found on Mary’s body. In other words, ‘John molested Mary’ allows us to predict that ‘there are biological traces of John on Mary’s body’ using the causal rule ‘person  $x$  molesting person  $y$  causes traces of  $x$  to be left on  $y$ ’s body’. Thus, the search for evidence is guided by the hypothetical scenarios considered: assuming  $S_1$  we should be able to find evidence of John’s traces on Mary’s body. This shows how causal scenarios can be used to both explain and predict evidence.

In inference to the best explanation, the objective is to consider the alternative scenarios and ultimately select the scenario that best explains the evidence. The theoretical question to be answered is: which scenario is the best explanation of the evidence? Pennington and Hastie (1993) provide several criteria for judging scenarios. The most important one, which is also standard in logical definitions of inference to the best explanation (Josephson and Josephson,



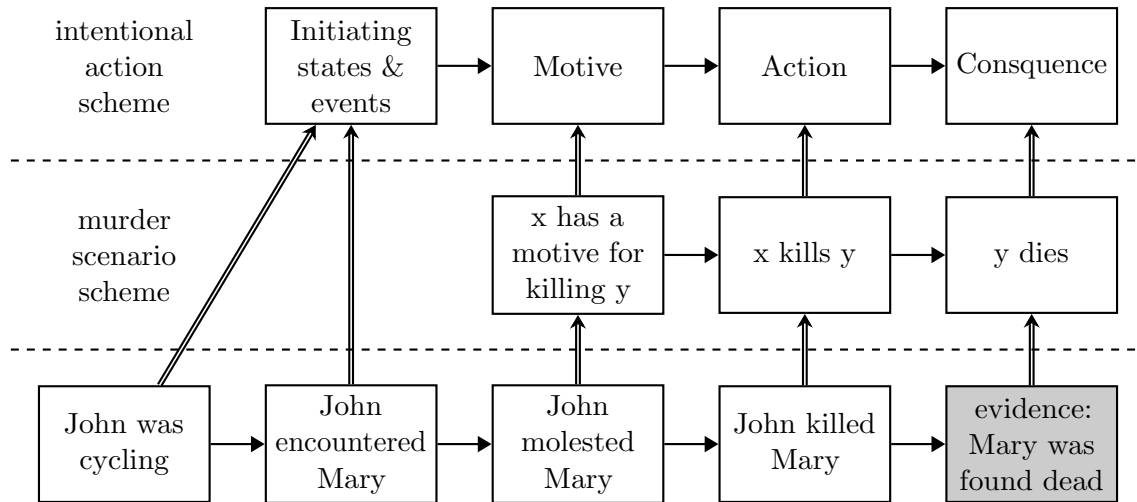


Figure 4: The scenario  $S_1$  as an instance of different scenario schemes

1996), is evidential coverage: how much of the evidence is covered by a particular explanation? For example, if we find evidence that the DNA of the traces left on Mary’s body matches John’s DNA, scenario  $S_1$  explains more evidence than  $S_2$ , which only explains the fact that Mary’s body was found.

In addition to looking at how well a scenario covers the evidence, it also makes sense to consider what Pennington and Hastie (1993) call the plausibility of a scenario irrespective of the evidence: does the scenario fit with our ideas about how things happen in the world? Whilst we would not want to convict a suspect purely on the basis of a plausible scenario which does not cover any of the evidence (the risk of choosing a “good scenario” over a “true scenario”, cf. Anderson et al., 2005; Bennett and Feldman, 1981), plausibility does play a big part in our reasoning. For example, the police will not seriously consider the scenario ‘Aliens killed Mary’ because this is highly implausible. Furthermore, elements which are implausible at first sight might warrant further investigation: why does John attack Mary when he encounters her? This is not normal behaviour for a man like John. Finally, judges or jurors are often also forced to fill gaps in the scenario using their own knowledge. For example, except in case of a confession, there is often no direct evidence for the fact that a killing was premeditated. In Dutch law, however, it is often accepted that the action was premeditated if it can be made plausible that, given the circumstances (i.e., the scenario) there was a moment in which the perpetrator could ‘calmly deliberate and consider’ his actions.

A notion related to the plausibility of scenarios is that of scenario schemes—also called story schemes (Bex, 2011) or scripts (Schank and Abelson, 1977)—stereotypical patterns that serve as a scheme for particular scenarios. Scenario schemes can be used to answer the question whether a scenario contains all its elements, and can hence be used to establish what we refer to as a scenario’s global coherence. For example, Pennington and Hastie (1993) use a general scenario scheme for intentional action: given some *initiating events and states of affairs*, a *motive* may lead to an *action* with certain *consequences*. More specific scenario schemes may be instances of such a generic scheme: a murder, for example, is a specific type of intentional action, where the *action* involves one person killing another and the *consequence* is that the victim dies. In Figure 4, the scenario  $S_1$  is rendered together with two scenario schemes, one for intentional actions and one for murder. The double arrows indicate abstraction relations. In the figure, the most abstract scheme is the intentional action scenario scheme; the murder scheme is a specialisation of this more general scheme, and the scenario  $S_1$  is an instance of both the murder

scheme and the intentional action scheme.

Whilst the plausibility of the individual causal generalisations also play a significant part in causal reasoning, scenario schemes are used for capturing the global coherence (Section 1) of scenarios. In order to determine whether a scenario is plausible and coherent, we can see whether it fits with well-known scenario schemes or whether any elements are missing. For example, a murder scenario with a missing motive is incomplete and therefore less coherent: a scenario where John suddenly kills Mary without molesting her, for example, is less coherent and plausible than  $S_1$ , because there is no real motive for the murder.

### 2.3 Probabilities

Influenced by the rise of DNA profiling and by some high profile miscarriages of justice, probabilistic approaches to reasoning with evidence remain a focus of study (Dawid et al., 2011; Fenton, 2011). Proposals and applications go back to the early days of forensic science (Taroni et al., 1998), and become modernised by making connections to computational modeling methods such as Bayesian networks (Taroni et al., 2006; Hepler et al., 2007; Fenton et al., 2013). The role of probabilistic reasoning remains an issue of debate, cf. also the recent discussion following the UK Court of Appeal decision to restrict the use of Bayes' theorem in courts to cases with a solid statistical foundation such as DNA (see the 2012 special issue of Law, Probability and Risk on the *R v T* case; Vol. 4, No. 2).

The use of probabilities can be illustrated in our example case with what we know about the DNA profile of the blood trace, found on the victims coat. It was estimated that the probability that the DNA profile of a random male matches the DNA profile of that blood trace was about 1 in 1500 billion billion, i.e.,  $0.66 \cdot 10^{-21}$ . Let us assume that there is a population of 8000 other males than John that could have murdered Mary (say, the local population), and that each of the 8001 males considered has equal probability of being the source of the DNA. As a result, the prior probability of one of these males to be the murderer is assumed to be  $1/8001$ . Often, assumptions need to be made about probabilities—resulting in subjective probabilities—in order to be able to perform the relevant probabilistic computations. We here use the number 8000, as that was (roughly) the number of men that were asked for a DNA sample, because they were living in the area, and were within reasonable age limits.

Let  $H_1$  be ‘John is the source of the DNA’ and let  $H_2$  be ‘Someone else is the source’, where  $H_1$  and  $H_2$  are each other’s negation. Writing  $E$  for ‘A DNA match was found’, we have the following three values:

$$\begin{aligned} P(H_1) &= 1/8001 \\ P(E|H_2) &= 0.66 \cdot 10^{-21} \\ P(E|H_1) &= 1 \end{aligned}$$

The third probability indicates that it is certain that a DNA match with John’s profile is found when John is the source of the DNA. The probability calculus allows the calculation of many other probabilities and conditional probabilities using the variables  $H_1$  and  $E$ , including the joint probability distribution over these variables, i.e., the values  $P(H_1 \wedge E)$ ,  $P(H_1 \wedge \neg E)$ ,  $P(\neg H_1 \wedge E)$  (equal to  $P(H_2 \wedge E)$ ) and  $P(\neg H_1 \wedge \neg E)$  (equal to  $P(H_2 \wedge \neg E)$ ).

Using these numbers, it is for instance possible to compute the especially relevant probability  $P(H_1|E)$  that John is the source given the match. It is this number that is often confused with the number  $P(E|H_1)$ , the probability of finding a match given that John is the source.  $P(H_1|E)$  can be computed using Bayes’ theorem:

$$P(H_1|E) = P(E|H_1) \cdot P(H_1) / P(E).$$

Of the three numbers on the right-hand side of this equation,  $P(E|H_1)$  and  $P(H_1)$  are given, so in order to apply the theorem we must first compute  $P(E)$ :

$$\begin{aligned} P(E) &= P(H_1 \wedge E) + P(\neg H_1 \wedge E) \\ &= P(E|H_1)P(H_1) + P(E|H_2)P(H_2) \\ &= 1 \cdot 1/8001 + 0.66 \cdot 10^{-21} \cdot 8000/8001 \\ &= (1 + 8000 \cdot 0.66 \cdot 10^{-21})/8001 \end{aligned}$$

Applying Bayes' theorem, we now find the probability  $P(H_1|E)$  that John is the source given the match:

$$P(H_1|E) = P(E|H_1) \cdot \frac{P(H_1)}{P(E)} = 1 \cdot \frac{1/8001}{(1 + 8000 \cdot 0.66 \cdot 10^{-21})/8001} = \frac{1}{1 + 8000 \cdot 0.66 \cdot 10^{-21}}$$

This is a number very close to the value 1, which indicates near certainty.

In probabilistic approaches to forensic science, Bayes's theorem in odds form (also known as Bayes's rule) plays an important role. This rule shows how the probability calculus governs the change of the odds of two complementary hypotheses before and after new evidence is found:

$$\frac{P(H_1|E)}{P(H_2|E)} = \frac{P(E|H_1)}{P(E|H_2)} \cdot \frac{P(H_1)}{P(H_2)}$$

The so-called posterior odds  $P(H_1|E)/P(H_2|E)$  on the left-hand side of the formula are found by multiplying the prior odds  $P(H_1)/P(H_2)$  with the likelihood ratio  $P(E|H_1)/P(E|H_2)$ . A high likelihood ratio indicates evidence that strongly distinguishes the two hypotheses. In our example, the likelihood ratio  $P(E|H_1)/P(E|H_2)$  is equal to  $1/0.66 \cdot 10^{-21} = 1.5 \cdot 10^{21}$ . Not surprisingly, this high number indicates that a DNA match is strongly distinguishing. Finding the match, the prior odds  $P(H_1)/P(H_2)$ , equal to  $1/8000$ , can be 'updated' to the posterior odds  $P(H_1|E)/P(H_2|E)$ , equal to  $1.5 \cdot 10^{21}/8000$ .

A probability function can be represented as a Bayesian network. A Bayesian network consists of an acyclic directed graph with the variables of the probability function as nodes. Each node has an associated probability table, specifying the probabilities of the node's values conditioned on all value combinations of the node's parents. Figure 5 shows the graph of a Bayesian network for variables  $H_1$ ,  $H_2$  and  $E$  as discussed above, and Table 2 the associated conditional probability tables for each node. A choice was made to model  $H_1$  and  $H_2$  as separate boolean nodes in the graphical structure, rather than as values of one single node. This is because in Section 4, we intend to model alternative scenarios in a Bayesian network, in which case the structure of a network will need to accommodate for more elaborate hypotheses (namely, full scenarios) which are not necessarily strictly each other's negation.

In the probability tables for this network (see Table 2) the number  $8000/8001$  in the first table is the probability  $P(H_2)$  that John is not the source. The number 0, bottom right in the second table, is the probability  $P(H_2|H_1)$  that someone else is the source given that John is the source. The number  $0.66 \cdot 10^{-21}$ , bottom left in the third table, indicates the probability of finding a DNA match, given that someone else is the source and John is not the source, i.e.,  $P(E|H_2 \wedge \neg H_1)$ . Since in this example  $H_1$  and  $H_2$  are each other's negation, this number is equal to  $P(E|\neg H_1)$  and to  $P(E|H_2)$ . Some values in the third table concern combinations of parents that cannot occur. We have entered the probability of the outcome given those situations as 0.5 but this choice is arbitrary since the values are irrelevant.

From a Bayesian network, any prior or posterior probability of interest can be computed. For example, the probability that John is the source given the match can be found by entering the evidence 'DNA match = true' into the network. Various software tools for working with Bayesian

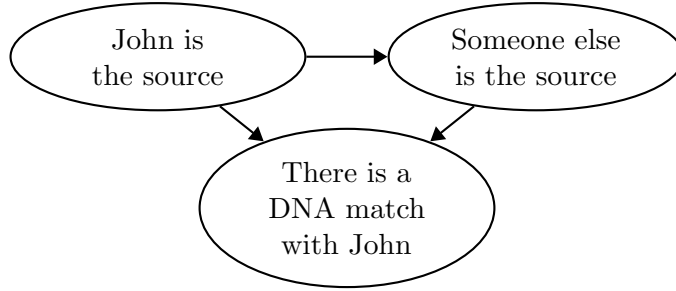


Figure 5: A Bayesian network structure with dependency relations

**John is the source**

John is the source = false	8000/8001
John is the source = true	1/8001

**Someone else is the source**

John is the source	false	true
Someone else is the source = false	0	1
Someone else is the source = true	1	0

**DNA match**

John is the source	false		true	
	Someone else = false	Someone else = true	Someone else = false	Someone else = true
DNA match = false	0.5*	$1 - 0.66 \cdot 10^{-21}$	0	0.5*
DNA match = true	0.5*	$0.66 \cdot 10^{-21}$	1	0.5*

Table 2: Conditional probability tables for the Bayesian network in Figure 5. All numbers have to be specified in a valid BN model, but the ones marked (\*) are irrelevant since the corresponding situations cannot be reached.

networks exists, such as GeNIe ([dslpitt.org/genie](http://dslpitt.org/genie)) or SamIam ([reasoning.cs.ucla.edu/samiam](http://reasoning.cs.ucla.edu/samiam)). Observing or instantiating a node to have a certain value (such as ‘DNA match = true’) will produce updated probabilities throughout the network, for instance an updated probability that John is the source, given the DNA match:  $P(\text{John is the source} = \text{true} \mid \text{DNA match} = \text{true})$ .

In a Bayesian network structure, the arrows contain information on the (in-)dependencies in the model. From the graph, it can be read whether there is possibly an influence between two variables  $A$  and  $B$ . However, such an influence can change as a result of instantiating nodes in the network. In Bayesian network terminology, d-connectedness and d-separation are the terms used to express whether there is a possible influence between nodes  $A$  and  $B$ . Whether nodes are d-connected or d-separated depends on whether there is an active chain between these nodes. Variables are d-connected when there is an active chain and when there is no active chain, the variables are d-separated.

Suppose three variables  $A$ ,  $B$  and  $C$  are connected via a serial connection:  $A \rightarrow C \rightarrow B$ . When  $C$  has not been observed, influence can pass from  $A$  to  $B$  via  $C$ , but also from  $B$  to  $A$  via  $C$ . This is an active chain, and  $A$  and  $B$  are d-connected. However, as soon as  $C$  is observed, the chain is blocked and when no other active chains remain,  $A$  and  $B$  are d-separated. A similar situation occurs when  $A$  and  $B$  are connected via  $C$  with a diverging connection:  $A \leftarrow C \rightarrow B$ . Again, this is an active chain as long as  $C$  has not been observed. As soon as  $C$  is instantiated,

the chain becomes blocked and when there are no other active chains between  $A$  and  $B$ , they are d-separated. A special situation is when  $A$  and  $B$  are connected via  $C$  with a converging connection:  $A \rightarrow C \leftarrow B$ . This is also called a head-to-head connection. As opposed to the previous situations, this chain is inactive as long as  $C$  has not been observed. When there are no other active chains between  $A$  and  $B$ , this means they are d-separated as long as  $C$  has not been observed. As soon as  $C$  or a descendant of  $C$  is instantiated, the chain becomes active and  $A$  and  $B$  are d-connected.

As an example of a converging connection, consider a structure with only three nodes in which  $A$  and  $B$  are alternative causes for a shared effect  $C$ . For example, let  $A$ ,  $B$  and  $C$  be as follows:  $A$ : it has rained,  $B$ : the sprinklers were on and  $C$ : the grass is wet. As long as  $C$  (wet grass) has not been observed, the two causes are d-separated and they have no influence on each other. When  $C$  is observed (the grass is wet), the two causes become d-connected, which can be understood as follows: knowing about one cause (sprinklers were on) means that the other cause (it has rained) becomes a less likely cause for this effect. Even though Bayesian networks need not be causally interpretable in general, such an effect between parents of a common child is referred to as an inter-causal interaction. This particular example is a very common type of inter-causal interaction called explaining away.

### 3 Connecting arguments and scenarios: a hybrid theory

Our thinking about the connections between different approaches to evidential reasoning started by comparing and connecting scenarios and arguments. In the research on legal theory and legal philosophy, there seemed to be two clear, competing approaches. The first is the “neo Wigmorean” approach (Anderson et al., 2005; Bex et al., 2003), in which evidential argument-trees for the possible conclusions in the case are constructed. The second is the narrative approach (Pennington and Hastie, 1993; Wagenaar et al., 1993; Josephson, 2002), where competing scenarios about “what happened” are constructed and compared. In the project “Making Sense of Evidence”, which ran from 2005 to 2009, some of the current authors tried to marry these two approaches in a formal hybrid theory of scenarios and arguments (Bex, 2014, 2011; Bex et al., 2010).

#### 3.1 On arguments and scenarios

Both the argumentative approach and the scenario-based approach can be separately applied to a case, and each of the two has their own advantages and disadvantages, as was also shown in Table 1. The argumentative approach is positioned in a formal dialectical framework (Dung, 1995) for adversarial reasoning and it is expressive enough to capture the different aspects of evidential reasoning (Bex et al., 2003). Furthermore, it has been argued that when given, for example, a witness testimony it is most natural for people to use evidential rules to infer a conclusion from this testimony (van den Braak et al., 2008).

The scenario approach captures the causal elements of a case (e.g., when talking about the cause of death, or when predicting which kinds of traces may have been left behind at the scene of the crime), and it also provides an overview of “what happened” in relation to the available evidence, thus making it suited for judging the global coherence of a case. While a standard formalisation of scenario-based reasoning is perhaps lacking, this type of reasoning can be captured by logical formalisations of model-based abductive reasoning (see, for instance, Josephson, 2002). Such formalisations use causal rules to model the scenarios, which are then compared on basic criteria such as the minimum number of assumptions. Because these frameworks were originally intended for automatic diagnosis within bounded and pre-defined domains, they are less suited to modelling more open-ended and complex large criminal cases

(Bex, 2011; Prakken and Renooij, 2001). For example, in purely scenario-based approaches, it is not possible to talk about the scenarios themselves; while in argumentation one can give a reason for an inference warrant, it is impossible to give a reason for a causal relation in a scenario.

Because both the scenario-based and the argument-based approach have their own advantages and disadvantages, a combination of the two seems to be an intuitive and analytically useful perspective. Hence, in (Bex et al., 2010; Bex, 2011) a *hybrid theory* of arguments and scenarios is proposed.

### 3.2 A formal hybrid theory

In the hybrid theory, causal-abductive scenario-based reasoning is combined with a general argumentation framework for evidential reasoning. Scenarios and sub-scenarios are used to explain ‘what happened’, and arguments are used to support or attack these scenarios with evidence. Furthermore, arguments also allow us to draw further (legal) conclusions from scenarios. Below, we will briefly discuss the formal hybrid theory (Bex, 2014, 2011; Bex et al., 2010) by means of the example in Figure 6, which shows the two scenarios  $S_1$  and  $S_2$  from Figure 3 and the arguments supporting and attacking them.

The hybrid theory consists of a set of evidence  $\mathbf{E}$ , a set of hypotheses  $\mathbf{H}$  and a set of inference rules  $\mathcal{R}$  of the form  $r_i : p_1 \ \& \ \dots \ \& \ p_n \Rightarrow_{C/E} q$ , where  $r_i$  is the name of the rule,  $\Rightarrow_C$  indicates a causal rule and  $\Rightarrow_E$  an evidential rule. Scenarios can then be built by assuming some hypotheses  $H \subseteq \mathbf{H}$  and consecutively applying causal inference rules to infer evidence  $E \subseteq \mathbf{E}$ . Arguments can be built by taking evidence  $E \subseteq \mathbf{E}$  and consecutively applying evidential inference rules.

As an example of scenarios, consider the evidence  $E = \{\text{Mary was found dead}\}$  and the following causal rules.

- $r_1^c$  person  $p$  was cycling  $\Rightarrow_C$   $p$  encountered Mary
- $r_2^c$  person  $p$  encountered Mary  $\Rightarrow_C$   $p$  molested Mary
- $r_3^c$  person  $p$  molested Mary  $\Rightarrow_C$   $p$  killed Mary
- $r_4^c$  person  $p$  killed Mary  $\Rightarrow_C$  Mary was found dead

Note how these rules are specific in the way in which they force any scenario based on them to be about Mary. This is not a problem, as the identity of the victim and what happened to her were not an issue in this case. The issue was exactly who the perpetrator (person  $p$ ) was: if we hypothesise that  $H = \{\text{John was cycling}\}$  we can infer the scenario  $S_1$  to explain Mary was found dead, and if we hypothesise that  $H = \{\text{AS was cycling}\}$  we can infer  $S_2$  which also explains Mary was found dead.

As an example of arguments, consider the evidence  $E = \{\text{DNA match John, no DNA match AS}\}$  and the following evidential rules.

- $r_1^e$  DNA match John  $\Rightarrow_E$  John is source of traces on Mary’s body
- $r_2^e$  John is source of traces on Mary’s body  $\Rightarrow_E$  John molested Mary
- $r_3^e$  no DNA match AS  $\Rightarrow_E$  AS is not source of traces on Mary’s body
- $r_4^e$  AS is not source of traces on Mary’s body  $\Rightarrow_E$   $\neg$ AS molested Mary

Now, if we take DNA match John we can infer the conclusion John molested Mary by successively applying the rules  $r_1^e$  and  $r_2^e$ , which gives us argument  $A_2$  from Figure 6. Similarly, we can take the evidence no DNA match AS and infer  $\neg$ AS molested Mary by successively applying rules  $r_3^e$  and  $r_4^e$ , thus building  $A_3$  (Figure 6).

We can now define how arguments and scenarios can be combined. We say that an argument *supports* a scenario if the conclusion of the argument is an element of the scenario, and the argument itself is not defeated by another argument. For example, the conclusion John molested Mary of  $A_2$  is an element of  $S_1$ , so  $A_2$  supports  $S_1$ . An argument *attacks* a scenario

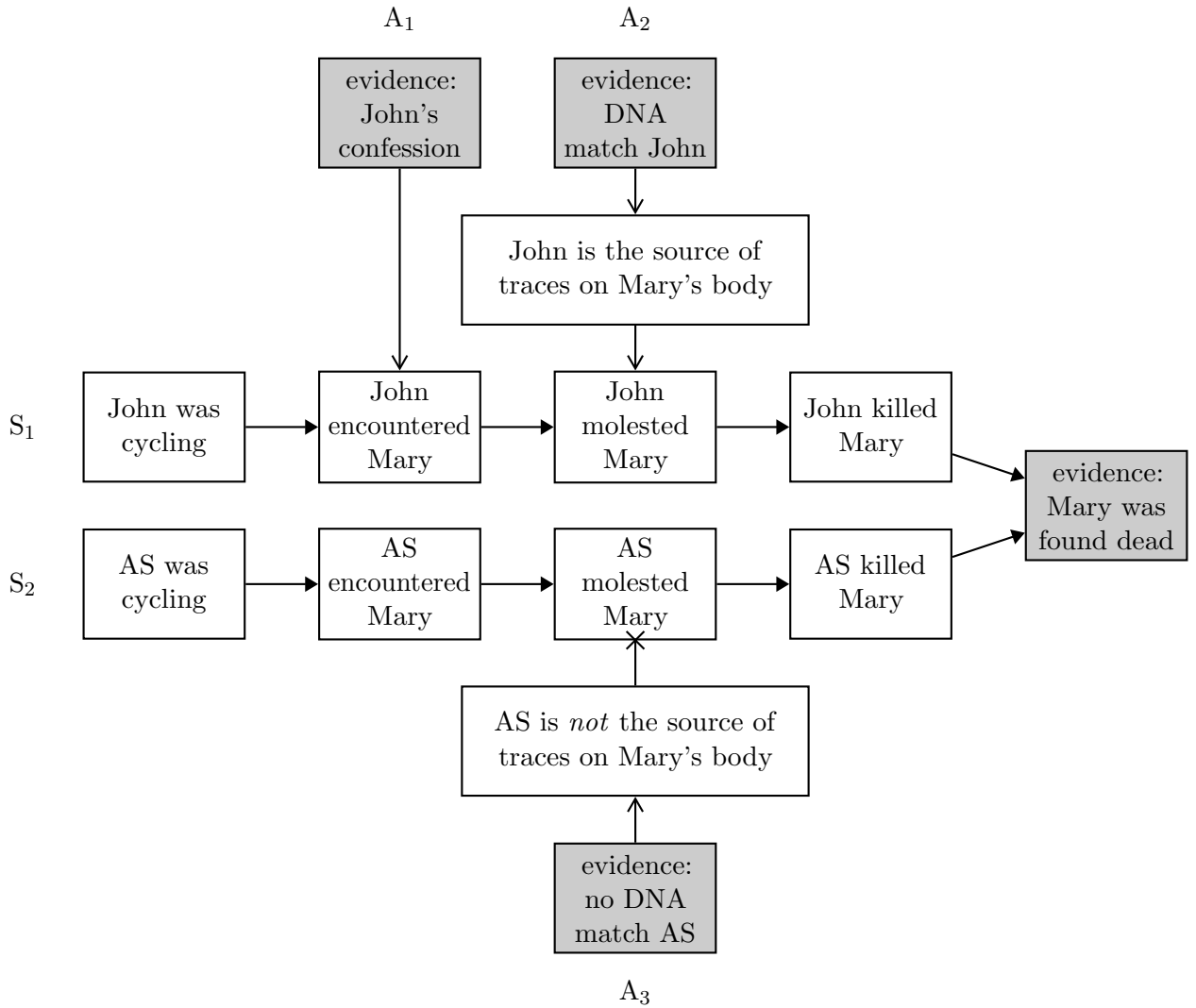


Figure 6: John’s scenario as a hybrid case. Arrows with open arrowheads stand for evidential inferences, and arrows with closed arrowheads stand for causal relations.

if the conclusion of the argument is the negation of an element of the scenario, and the argument itself is not defeated by another argument. For example, the conclusion  $\neg$ AS molested Mary of  $A_3$  is the negation of an element of  $S_2$ , so  $A_3$  attacks  $S_2$ . Note that it is also possible to support or attack (applications of) a causal rule  $r_i^c$  in a scenario by building an argument for the conclusion  $r_i^c$  or  $\neg r_i^c$ , respectively.

### 3.2.1 Scenario schemes and scenario hierarchies

While explicit causal rules play an important part in the hybrid theory – they are used to infer the evidence from the hypotheses – recent work (Bex and Verheij, 2013) focuses more on global (‘holistic’) coherence of scenarios where causal coherence is not explicitly represented but rather left implicit. To this end, scenario schemes are defined that can be used for the construction of hypothetical stories. As an example, take the scheme for murder  $ss_1$  which was also mentioned in Figure 4.

$ss_1$  [x had motive for killing y, x killed y, y was found dead]

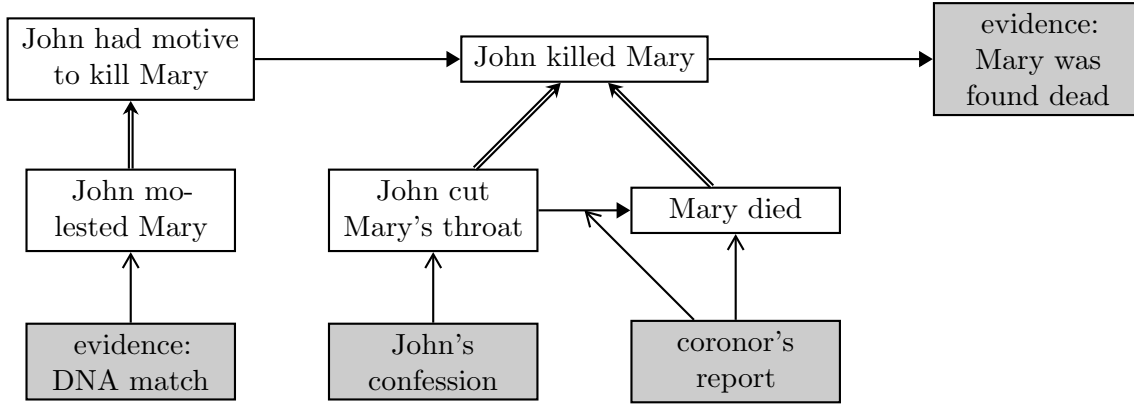


Figure 7: General and specific scenario elements

Now, if we hypothesise that  $H = \{\text{John had motive for killing Mary}\}$  we can instantiate  $x$  with **John** and  $y$  with **Mary** and hence we can explain **Mary was found dead**. Even though the causal links are left implicit, we still have a fairly coherent (if overly generic) scenario.

Recall from section 2.2 that it is possible to use abstraction links to connect scenarios to more abstract schemes (the double arrows in Figure 4). These same abstraction links, which are of the form  $r_i : p_1 \ \& \ \dots \ \& \ p_n \Rightarrow_A q$ , can also be used to connect scenarios at different levels of abstraction (Console and Theseider Dupré, 1994). In this way, elements of a scenario can be ‘unfolded’ into a more specific sub-scenario.<sup>5</sup> Take, for example, the following abstraction and causal rules.

$$\begin{aligned}
 r_1^a \quad & \text{John molested Mary} \Rightarrow_A \text{John had motive to kill Mary} \\
 r_2^a \quad & (\text{John cut Mary's throat} \ \& \ \text{Mary died}) \Rightarrow_A \text{John killed Mary} \\
 r_5^c \quad & \text{John cut Mary's throat} \Rightarrow_C \text{Mary died}
 \end{aligned}$$

Using these rules we can infer, for example, the more general **John had motive to kill Mary** from **John molested Mary**, and the general **John killed Mary** from the specific elements **John cut Mary's throat** and **Mary died**. In Figure 7, these relations between scenarios and specific sub-scenarios are indicated. If we now assume that  $H = \{\text{John molested Mary, John cut Mary's throat}\}$ , we can infer the evidence **Mary was found dead** through a combination of causal and abstraction rules. Notice how the evidential arguments support the relevant sub-scenarios, including the application of the causal rule  $r_5^c$ .

### 3.2.2 Comparing scenarios in a case

Given alternative scenarios such as  $S_1$  and  $S_2$ , the question is now how to compare them. In Bex (2011), a number of criteria for comparing scenarios is given. An important one is *evidential coverage*, the portion of the evidence in a case that supports the scenario (see also section 2.2). In the example,  $S_2$  has an evidential coverage of  $1/4$ , as there are 4 pieces of evidence and only the evidence ‘Mary is found dead’ supports  $S_2$ , and  $S_1$ ’s evidential coverage is  $3/4$ , as three pieces of evidence support  $S_1$ . Related to evidential coverage is *evidential contradiction*, which is the portion of the arguments based on evidence that contradict a scenario. In the example,  $S_1$  has an evidential contradiction of 0 (no argument based on evidence attacks  $S_1$ ) and  $S_2$  has an evidential contradiction of  $1/4$ , as the evidence about the DNA that did not match attacks  $S_2$ . Note that the two evidential criteria do not give an absolute measure of how good or strong a scenario is. It is for example possible that even though one scenario explains a lot of evidence,

<sup>5</sup>The idea of ‘unfolding’ a scenario into sub-scenarios was coined by Vlek et al. (2014a), see section 4.2.



it does not cover a crucial piece of evidence (e.g., very strong DNA evidence). However, the coverage and contradiction can be used as relative measures to compare scenarios and guide the search for further evidence; if a plausible position has low evidential coverage it might make sense to search for evidence that supports the position.

Another way to compare scenarios is by looking at their coherence irrespective of the evidence in a case. In other words, is the scenario plausible given our general knowledge about the world? Here, scenario schemes play an important part, as we have to look whether the scenario fits a particular scheme, and whether that scheme is a plausible generalisation of how things normally happen in the world. In the example case, we could say that the asylum seeker scenario is *prima facie* more plausible than the scenario about John: John is known as a family man in the village. The asylum seekers came from conflict areas and might be traumatised, causing them to act out violently. Whilst this is a valid way of reasoning, it also demonstrates the danger of scenarios and scenario schemes, because they often appeal to certain stereotypes. We should therefore be careful with drawing conclusions from scenarios that have no evidence to back them up.

### 3.2.3 Scenarios and arguments: two sides of the same coin

Whilst developing the hybrid theory, we saw many similarities to argumentative thinking in the work by Wagenaar et al. (1993) and their anchored narratives theory, which was nevertheless explicitly promoted as a pure scenario approach. Moreover, it was obvious that practitioners in the field (investigators, fact-finders, lawyers) seemed to naturally combine argumentative and scenario elements in their work. Finally, causal and evidential reasoning are closely entwined: if we have a causal rule  $c$  causes  $e$  and  $c$  is the usual cause of  $e$ , then we will usually also accept that  $e$  is evidence for  $c$  (Pearl, 1988). Thus, it seems that causal scenarios and evidential arguments are not two separate approaches but rather two sides of the same coin.

As an example, consider the different ways in which the inferences and attacks surrounding DNA evidence can be captured. One way to do this is to use an evidential rule, like in Figure 8. This argument, that the match of the DNA of the traces with John’s DNA is evidence for the fact that John is the source of the traces, is undercut by stating that the sample was tainted. However, we can also say that the DNA match was *caused* by John being the source of the traces, changing the reasoning from evidential argumentation to causal scenario-based reasoning. The attack is then captured as an alternative explanation of the DNA match evidence: maybe the match was caused by the sample being tainted with John’s DNA in the lab. This shows that there is a clear link between alternative explanations and attacking arguments, a link which has recently been formalised in Bex (2014).

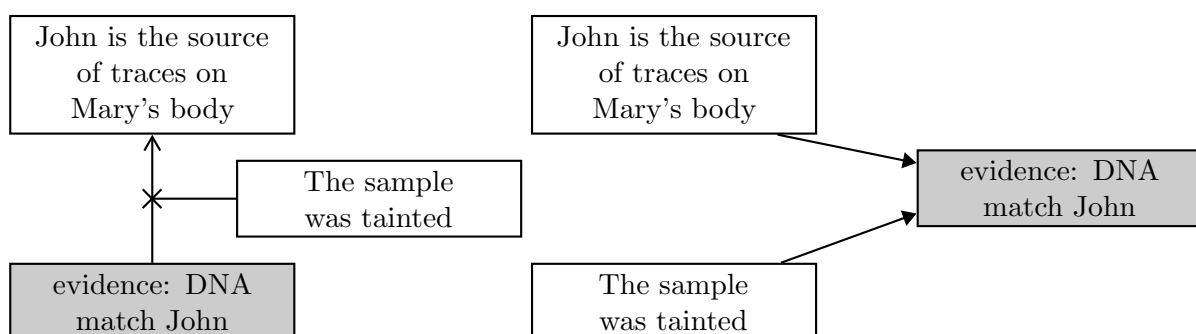


Figure 8: Reasoning about the DNA match evidence modelled as two attacking arguments or as two alternative (conflicting) scenarios

### 3.3 Strengths and limitations

In sum, we have proposed a formal model connecting arguments and scenarios in evidential reasoning. Strengths and limitations of the proposal include the following:

#### Strengths

1. The proposal shows how reasoning with arguments can be formally combined with reasoning to the best explanatory scenario.
2. The proposal provides a formal analysis of reasoning about evidential and causal rules.
3. The use of scenario schemes emphasises the role of the global coherence of scenarios.

#### Limitations

1. In the proposal, there is no modelling of degrees of uncertainty.
2. The formalisation is based on argumentation formalisms and these are (as yet) not standardised nor well connected to standard theories, such as classical logic and standard probability theory.

## 4 Connecting scenarios and probabilities: embedding scenarios in Bayesian networks

The combination of arguments and scenarios in Section 3 made it possible to construct arguments to support scenarios, and to reason about the internal coherence of scenarios. However, it was impossible to reason about degrees of uncertainty in that approach, whereas due to the importance of DNA evidence such reasoning is needed. For instance, the method from Section 3 does not distinguish between strong evidence and weaker evidence, when comparing alternative scenarios. In this section, the connection between scenarios and probabilities is discussed. The probabilistic framework of Bayesian networks allows for modelling degrees of uncertainty concerning the evidence, which makes it possible to incorporate the strength of evidence in a decision. In what follows, we will discuss how a scenario can be modelled in a Bayesian network such that the key properties of a scenario are represented probabilistically.

### 4.1 On scenarios and Bayesian networks

A main feature that distinguishes a scenario from any other collection of events is a scenario's *coherence*. The elements of a scenario together form a coherent whole. This was described by Pennington and Hastie (1993) as a scenario 'having all of its parts'. As a consequence of coherence, scenarios can be used to reason about hypothetical events for which there is no direct evidence (Tillers, 2005). This needs to be captured probabilistically in order to model scenarios in a Bayesian network.

Consider the scenario about John molesting and killing Mary. Suppose that there is evidence to support that John was the molester, but no evidence to support that he was the killer (in the real case this was more or less the situation after the DNA screening but before John confessed). In this situation, John killing Mary is an *evidential gap* in the scenario. But due to the scenario's coherence, the evidence for the molestation also increases our belief in the event that John killed Mary. Despite the evidential gap, this scenario can still be used to reason about the killing.

When reasoning with evidential gaps as described above, our degree of belief in one element of the scenario (namely, that of John killing Mary) increased because other elements of the

scenario became more believable (as a result of the supporting evidence). This is called *transfer of evidential support*, and this is what we aim to capture in a Bayesian network model of a scenario: there is an influence between all elements of a scenario, such that when one element becomes more probable given the evidence, other elements become more probable as well.

To capture scenarios and their coherence in a Bayesian network, we propose the use of *idioms*. An idiom is a general structure that can be used as a building block in a Bayesian network, simplifying the task of constructing a network structure. The idea of such recurrent substructures for building legal Bayesian networks was proposed by Hepler et al. (2007) and later extended by Fenton et al. (2013), who compiled a list of *legal idioms*. The concept is similar to the concepts of argumentation schemes and scenario schemes (see Section 3), in which typical patterns of arguments and scenarios, respectively, are modelled. However, idioms are less context-dependent than argument and scenario schemes, and can be used as building blocks throughout various cases.

In Section 4.2, two idioms are proposed: the *scenario idiom* for capturing the coherence of a scenario, and the *subscenario idiom* for capturing smaller subscenarios within the main scenario, which have their own internal coherence. Both idioms were previously described by Vlek et al. (2014a) (among other idioms), as part of a design method for an incremental construction of a Bayesian network using several alternative scenarios as a basis.

## 4.2 The scenario idiom and the subscenario idiom

The scenario idiom and the subscenario idiom capture coherence of a scenario, possibly with subscenarios. As discussed in Section 4.1, what we specifically want to capture in our models is the transfer of evidential support, which means that when one element of a scenario becomes more probable, the probabilities of other elements also increase. In the scenario idiom (see Figure 9(a)), all elements of the scenario are modelled as separate boolean nodes (element nodes  $P_1, P_2, \dots$ ), and arrows between elements of the scenario are drawn whenever connections are present within the scenario (shown as dashed arrows in the figure). To capture the coherence of a scenario, an additional boolean node called the ‘scenario node’ is placed at the top, and arrows are drawn from the scenario node to each of the element nodes. Considering that the scenario node itself is never observed, these arrows ensure an influence between elements of the scenario (needed to capture the transfer of evidential support as discussed in the previous section) via

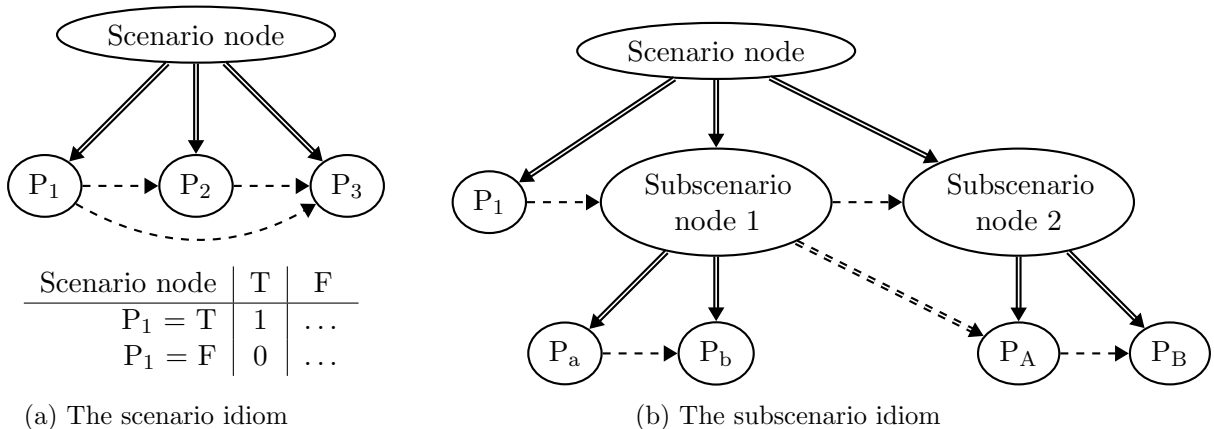


Figure 9: The scenario idiom (left) and the subscenario idiom (right). Double arrows signify that the underlying probabilities are partially fixed as shown in the table. Dashed arrows show some possible connections.

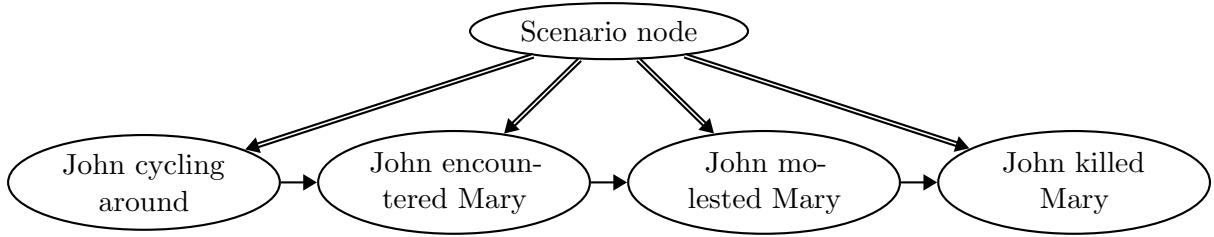


Figure 10: The scenario about John

the scenario node.

The scenario node represents the scenario as a whole. The probability table of the scenario node thus requires a prior probability for the scenario being true. Furthermore, due to the nature of the scenario node, representing the scenario as a whole, there is a special relation between the scenario node and each element node, signified by double arrows in Figure 9. The intuition is the following: whenever a scenario as a whole, represented by the scenario node, is true, each element of the scenario must be true. Because of this, some numbers in the probability table of each element node are fixed: an example of a probability table for an element node is shown in the table in Figure 9(a). With these probabilities, the transfer of evidential support is captured, since (in the absence of other influences) an increased belief in one element of a scenario will lead to an increased belief in the scenario node, which in turn yields an increased belief of all other element nodes.

With the scenario idiom, the scenario about John can be modelled as shown in Figure 10. Due to the structure of the scenario idiom and the probabilities as specified in the table of Figure 9(a), the transfer of evidential support is guaranteed. This means that, as was described in Section 4.1, once there is evidence for John molesting Mary (which will be modelled as a separate node connected to the node ‘John molested Mary’), the probability of John killing Mary will also increase.

The subscenario idiom builds upon the same ideas as the scenario idiom, but also captures the internal coherence of a subscenario. In order to model a coherent subscenario within a scenario, a subscenario node is used to represent the subscenario as a whole, and arrows with probabilities fixed similarly are drawn from the subscenario node to all elements in that subscenario as shown in Figure 9(b). Again, transfer of evidential support within a subscenario is guaranteed.

With the scenario idiom and the subscenario idiom, it becomes possible to gradually construct a Bayesian network for a case (see Vlek et al. (2014a) for more on the design method). To construct a network, we rely on the concept of unfolding (which was already mentioned in Section 3.2.1). A scenario can be told at various levels of detail, and elements of a scenario can be unfolded into more specific subscenarios when needed. The construction process starts with an initial scenario such as the one from Figure 10. In order to include more information about the killing, the node ‘John killed Mary’ is *unfolded* to form a subscenario, as shown in Figure 11. The node itself now serves as a subscenario node, and the events of the subscenario are attached to that subscenario node. This process is repeated to gradually construct a Bayesian network with the required level of detail.

### 4.3 Strengths and limitations

To summarise, the approach described in this section has the following strengths and limitations:

#### Strengths

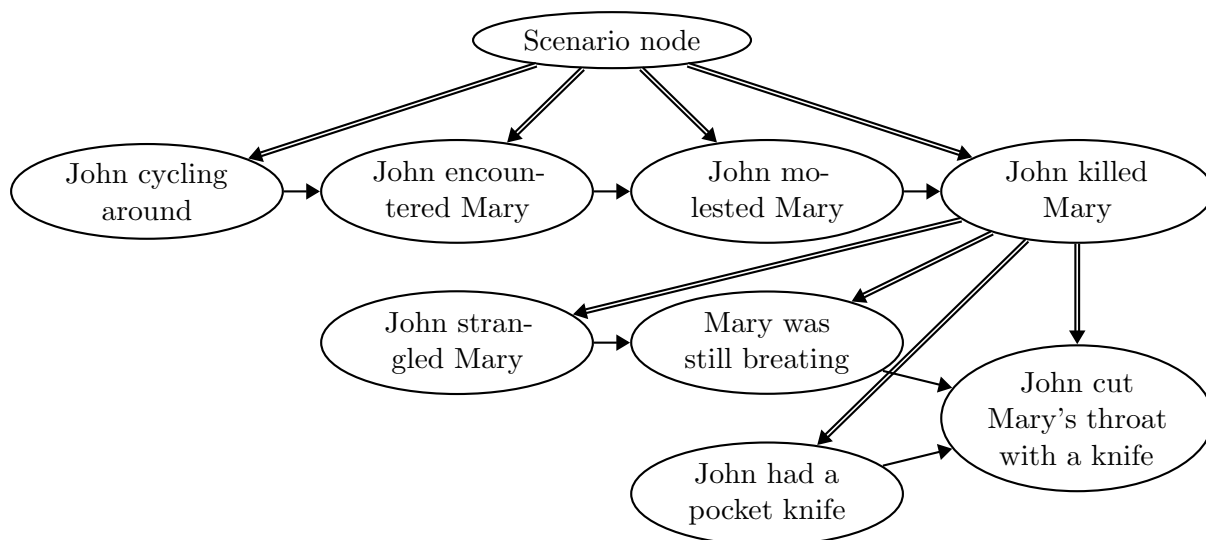


Figure 11: The event ‘John killed Mary’ was unfolded to a subscenario.

1. We have combined scenarios and their global coherence with degrees of uncertainty by showing how scenarios can be embedded in Bayesian networks, a prominent probabilistic modelling tool.
2. In the approach, we captured the concepts of coherence and transfer of evidential support probabilistically.
3. We have provided a probabilistic model of the unfolding of scenarios.

### Limitations

1. The approach inherits a standard criticism associated with Bayesian networks: since a Bayesian network is a model of a full probability function, more numbers are needed than are available, or can reasonably be estimated.
2. Bayesian network models including scenarios are large and complex, so explaining their meaning to fact-finders and forensic experts requires further study.

## 5 Connecting arguments and probabilities: extracting arguments from Bayesian networks

We have already seen argumentation and Bayesian networks in two different contexts now. Argumentation has been introduced as a method to support or attack events and causal links in a scenario-based model of evidence. These scenario models have been shown to be useful during the construction of Bayesian networks. There is, however, also a more direct connection between Bayesian and argumentative models of proof. We will proceed by describing some of the characteristics of both models and how these two formalisms can be used together. Specifically, we will show how arguments can be grounded using rules that can be extracted from a Bayesian network. Figure 12 shows a global outline of the approach. We make an automated translation from information in a Bayesian network to argumentation structures that can be used to support other arguments. The following method is based on Timmer et al. (2014).

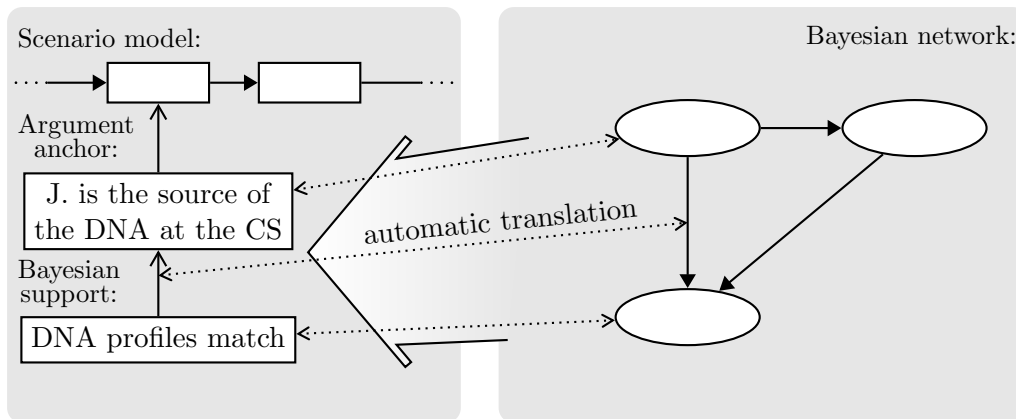


Figure 12: A high-level description of the Bayesian network-Argument translation approach.

## 5.1 On argumentation and Bayesian networks

To understand how we can extract arguments from a Bayesian network we must first identify what characteristics of probabilistic reasoning we would like to be able to capture.

Bayesian networks represent a joint probability distribution and as such can be a probabilistically accurate representation of the facts in a legal case. However, the structure of a Bayesian network is made to represent independence information (through d-separation) rather than inferential steps such as in many argumentative models. This mismatch in interpretation is what makes Bayesian network models less ideal for communication to legal experts such as lawyers and judges.

The directions of arrows in the Bayesian network convey some subtle information. They are easily misunderstood for causal relations, which they can be, but which is not the only possible interpretation. Without further information they just represent possible correlation. The directions of the edges contain information on the (in-)dependencies in the model.

To recall the example that was already introduced in Figure 5 and Table 2, remember that the Bayesian network modelled two perpetrator hypotheses and one piece of evidence—the DNA matching test. These two hypotheses can both individually cause a particular outcome of the DNA matching test. Because the two hypotheses are modelled—exactly for this reason—as parents of the evidence, finding one of them to be true explains the other away. If John is the source, then, likely, nobody else is and, vice versa, if another person is the source of the sample, John cannot be the source.

In argumentation links represent inference steps. The use of defeasible reasoning is particularly useful in combination with probability theory because statistical inferences are also not strict but merely suggest an elevated belief in some statements. In particular, the concept of undercutting has a striking resemblance to the concept of explaining away in Bayesian networks. An alternative explanation can provide a context in which the statistical inference is not applicable. In defeasible reasoning sometimes a rule that is in principle valid can not be applied to a premise, due to some exceptional circumstances. These circumstances are then called undercutters of the rule. In a probabilistic setting an explaining away provides a similar mechanism. This also resembles the way in which alternative explanations in scenarios are linked to attacking arguments as we discussed in Section 3.2.3.

## 5.2 Argument extraction

We now define an argumentation system (which is in fact a special case of the ASPIC+ argumentation framework (Modgil and Prakken, 2013) and very similar to the one used in the hybrid theory presented in Section 3) for Bayesian network argumentation. We use as a logical language ( $\mathcal{L}$ ) the set of all variable assignments  $V_i = v_{ij}$  in the Bayesian network model. A natural definition of negation follows from the fact that assignments for a node are mutually exclusive; all mutually exclusive assignments to a variable negate each other.

We extract defeasible rules from the Bayesian network by looking at a probabilistic measure of inferential strength. We enumerate candidate rules and assign strengths to them according to the so-called normalised likelihood. A number of different measures of strength have been introduced throughout the literature, and, while they often vary in exact numerical valuations of inferences, for many of these measures the ordering is proven to be the same (see Crupi et al. (2007) for an overview of these measures). Since we are going to use the measure of strength only to compare inference rules any of these will do.

**Definition 1** (strength (Timmer et al., 2014)). *A rule  $H_1, \dots, H_n \Rightarrow H$  has strength:*

$$\text{strength}(H_1, \dots, H_n \Rightarrow H \mid \mathbf{E}^*) = \frac{\text{P}(H \mid H_1, \dots, H_n \wedge \mathbf{E}^*)}{\text{P}(H \mid \mathbf{E}^*)}$$

*in which the evidential context  $\mathbf{E}^*$  is all the available evidence  $\mathbf{E}$  except those assigning a value to any of the variables from  $\{H, H_1, \dots, H_n\}$ .*

We construct a set of accepted rules  $\mathcal{R}_d$  of all rules with a strength greater than one. This is not an arbitrary threshold but a fundamental one. Since even rules with a strength slightly greater than one have a positive effect on the conclusion. This choice means that we accept every rule, however weak it may be. If the strength equals one, then the premises are independent of the conclusion and if the strength is below one the premises actually have a negative effect on the conclusion. In the latter case another rule with the opposite conclusion will automatically have a positive strength.

The evidential context  $\mathbf{E}^*$  is used to condition only on those evidence variables that do not occur in the premise or the conclusion of the rule under consideration. This is necessary because the rules have a counterfactual character. When some evidence is present, we have to speculate on what would have happened if it had not been the case, to be able to say something about the correlation of the evidence with other variables. In our running example, when the DNA evidence is observed, it becomes almost certain that John committed the crime. If we want to evaluate the strength of this evidence, we want to do this in the presence of all other evidence that was already there ( $\mathbf{E}$ ). However we do not want to condition the strength of the rule on the presence of the DNA evidence. If we did that, it would be as if we calculate the strength of the DNA evidence in a case where the DNA evidence is certain to match the suspect. Evidently, in such a case the evidence has no added value at all. Numerically we can also see that if we were to have the conclusion of the rule in  $\mathbf{E}^*$ , then both the numerator and the denominator become 1.0 exactly. Similarly, we can see that if one of  $H_1, \dots, H_n$  would be in  $\mathbf{E}^*$ , the numerator and denominator would become equal because they condition on exactly the same set.

We enumerate rules for every possible conclusion and with every possible set of premises with the restriction that premises must assign values to neighbours (parents and children) in the Bayesian network graph or to parents of children in the Bayesian network graph. The latter is necessary to capture the cases where inter-causal interactions occur. In that case we also require that at least one of the child-nodes is present in the set of premises for the rule, otherwise the head-to-head connection would not have been unblocked.

Defeasible rules can have exceptions. In a probabilistic setting, inferences can often be weakened (or invalidated altogether) by observing further evidence. Therefore, we identify undercutting variable assignments by checking if the measure of strength drops below one given any potentially undercutting variable assignment.

With the rules we can build an argumentation system. The knowledge base for this system will consist of all the observations in the Bayesian network. Using the set of rules we can build up arguments from that knowledge base. We iteratively apply the rules to the conclusions of other arguments, meanwhile making sure that every premise of the rule is fulfilled by the conclusion of another argument or by an observation from the knowledge base. We continue applying rules to the set of arguments until no more rules can be applied.

In the system described so far, we correctly represent inter-causal inferences in the argumentation system but we have not yet discussed how we can prevent the system from chaining rules from a parent to a child and then to another parent of that child. This would be the incorrect way to apply inferences because the inter-causal effect already models the relation between these variables and the strengths of the individual links have no bearing on the strength of the inter-causal interaction. This can be solved by either post-processing the arguments and filtering the ones out that have a reasoning step like this. A more elegant solution is to do it on the fly by maintaining some labels on the statements and in the rules in a way similar to Pearl’s CE-Logic (Pearl, 1988). Basically, what this does is to remember whether an inference along the direction of a Bayesian network edge has already been made and prevents the application of a rule against the direction of an edge in the Bayesian network to these statements.

### 5.3 Applications to the running example

To demonstrate the method we will show how it can be applied to our running example case. Figure 5 shows a Bayesian network of the probabilistic information that is given in the report by the court.

The approach described above can be completely automated. Applying this method to our example yields a number of probabilistically supported rules. For demonstration purposes we have displayed some of these rules in Table 3. We should immediately notice that some of the displayed rules have premises directly opposing the instantiations that we observed. These rules will be of no use although they are technically valid inferences. Instead of listing the rules, it is, of course, more informative to show how they can be combined into arguments. We have also automated that process and some of the interesting arguments are displayed in Figure 13. One of the drawbacks of the automated process of argument construction is that numerous arguments are constructed for conclusions that are not of interest to the user. For instance, the system will find (and therefore try to apply) rules with the DNA\_match node in the conclusion, simply because a system like this does not know what the variables of interest are.

Reasoning from the evidence node at the bottom upwards, we can see that the DNA match is a reason for both John being the source ( $A_2$ ) and someone else not being the source ( $A_1$ ). These two conclusions also support each other. Note that this results in distinct arguments with the same conclusion, because arguing that ‘the evidence suggests that john is the source, so no one else can be the source’ (as modelled in  $A_3$ ), is argumentatively different from arguing that ‘the evidence suggest that someone else is not the source’ (as in  $A_1$ ).

Also note that the single piece of probabilistic evidence that is modelled in the Bayesian network results in an argumentation model without attacks. If there is no contradicting information modelled by the evidence this is exactly what we expect. For the extracted rules we have also automatically identified possible undercutters but none of these can be supported by evidence so no attacking arguments can be constructed. An example of such an undercutter



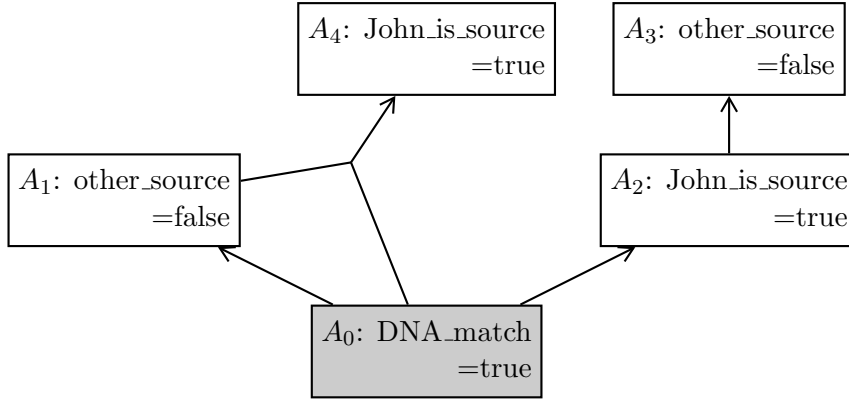


Figure 13: An argument graph resulting from the rules that we extracted from the Bayesian network.

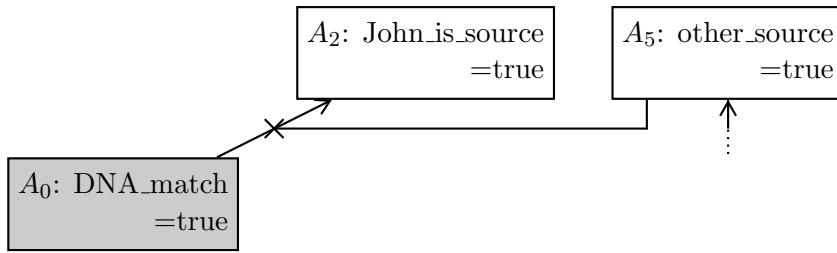


Figure 14: A hypothetical argument that applies a possible undercutter to Argument  $A_2$ .

is the assignment ‘other\_source=true’ which undercuts the rule ‘DNA\_match=true⇒John\_is\_source=true’. Figure 14 shows a hypothetical argument that harnesses this undercutter to attack argument  $A_2$  from Figure 13.

## 5.4 Strengths and limitations

In this section, we investigated connections between arguments and probabilities as normative frameworks in evidential reasoning. The approach includes these strengths and limitations:

### Strengths

premises	conclusion	strength
john_is_source=true	⇒ other_source=false	$\infty$
john_is_source=false	⇒ other_source=true	$\infty$
DNA_match=true	⇒ other_source=false	$1.20 * 10^{25}$
john_is_source=true	⇒ DNA_match=true	$\infty$
john_is_source=false	⇒ DNA_match=true	$\infty$
DNA_match=true, other_source=false	⇒ john_is_source=true	$\infty$
DNA_match=false, other_source=true	⇒ john_is_source=false	$\infty$
DNA_match=false	⇒ john_is_source=false	$\infty$
DNA_match=true	⇒ john_is_source=true	$1.20 * 10^{25}$
...	⇒ ...	...

Table 3: Some of the rules that were extracted from the example Bayesian network. Note that for practical purposes we have slightly abbreviated names of nodes.

1. We have shown how arguments can be extracted from Bayesian networks.
2. We have formally defined the strength of the rules and exceptions used to construct arguments.
3. The arguments extracted from a Bayesian network can help explain such networks, even when they are complex.

### Limitations

1. The developed argument extraction algorithms require computational resources that grow exponentially with the size of the network.
2. The arguments extracted from a network include many small variations, reducing their explanatory value.

## 6 Connecting arguments, scenarios and probabilities: arguments for and against scenarios in standard probability theory

In the previous sections, research on pairwise connections between the uses of arguments, scenarios and probabilities in reasoning with evidence has been discussed. Whereas sections 3 and 5 built on the argumentation formalism ASPIC+ (Prakken, 2010), and sections 4 and 5 on Bayesian networks (Jensen and Nielsen, 2007), the formal background of this section is standard probability theory and its underlying classical logic, following the proposal by Verheij (2014b), in connection with the formal discussion in Verheij (2014a).

The proposal reported on investigates a view on arguments to and from scenarios in the context of probability theory. The focus is on the arguments for and against the different, mutually incompatible scenarios available. The internal structure of arguments and scenarios, e.g., following argumentation and scenario schemes, is here not elaborated on. Arguments can have different strengths, measured using numbers that behave like conditional probabilities. It is not assumed that the strength of each and every argument is available. Some argument strengths can be established or sensibly estimated, but we accept that for many numbers there is no feasible way of determining their value. In this way, we can keep the constructive normative role of standard probability theory, without requiring that more numbers are available than can be reasonably expected.

### 6.1 Arguments to and from scenarios, with strengths as conditional probabilities

In the proposal, arguments have a strength that is measured as a conditional probability. Arguments can go from the evidence to a scenario, and from a scenario to the expectations based on the scenario. Different scenarios can be incompatible, and give rise to counterarguments for the arguments supporting them. The proposal to model arguments to and from scenarios in the context of probability theory is illustrated in Figure 15. In the figure, there is an argument from the combined evidence  $E$  to the hypothetical scenario  $H_1$  (with strength  $P(H_1|E)$ ), another from  $E$  to an incompatible scenario  $H_2$  (with strength  $P(H_2|E)$ ), and an argument from  $H_1$  to expectations based on the scenario. Since the scenarios  $H_1$  and  $H_2$  are incompatible, the arguments to these scenarios attack each other.

Characteristics of the proposal are the following:

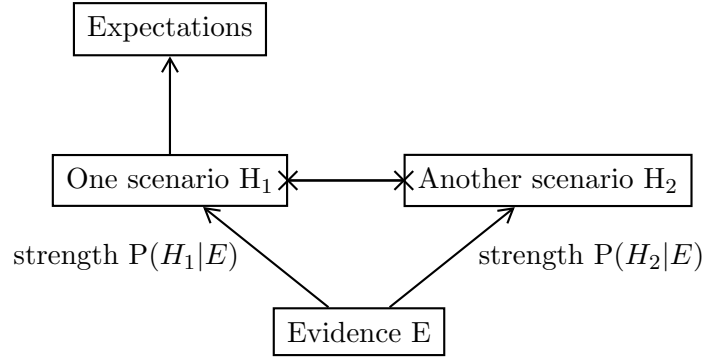


Figure 15: Arguments for incompatible scenarios, with their strengths (Verheij, 2014b)

1. Evidential reasoning is modelled as a process in which a model of the case in terms of the evidence, hypothetical scenarios and expectations is gradually developed. During the process, new evidence becomes available, and new hypothetical scenarios about what has happened are considered. Hypothetical scenarios are tested on the basis of expectations. If an expectation is contradicted by further evidence, the scenario is excluded.
2. The model of the evidence, hypothetical scenarios and expectations is developed within standard probability theory and its underlying classical logic. Typically, such a model does not specify a full probability function (as in a Bayesian network approach), since it is not assumed that all numbers are available, or can even be reasonably determined. In general, there can be many full probability functions that fit the model.
3. The aim of evidential reasoning is to develop a model about the case in which the established evidence leaves only one possible hypothetical scenario, while all alternative scenarios are impossible. That hypothetical scenario is then certain, given the evidence, according to the model of the case.

The model of the evidence, the hypothetical scenarios and expectations is developed in terms of probabilistic statements about the positions and reasons involved, where reasons are elementary, unstructured arguments that can be the building blocks of larger arguments. The strength of a reason is measured as a conditional probability, in contrast with contemporary approaches that follow Pollock’s work on defeasible argumentation, which uses an explicitly anti-probabilistic treatment of arguments and their strengths (Pollock, 1995, p. 99; 2010, p. 11; see Verheij, 2014a).

A reason is here a pair of sentences  $(\varphi, \psi)$ , where  $\varphi$  and  $\psi$  are sentences in a logical language. The strength of a reason  $(\varphi, \psi)$  is measured as the conditional probability  $P(\psi|\varphi)$ , where  $P$  is a function that obeys the properties of a standard probability function, hence only defined when  $P(\varphi) > 0$ .<sup>6</sup> When the strength of a reason is equal to 1, it is said to be conclusive: Then the reason’s conclusions  $\psi$  are certain given its premises  $\varphi$ . If the strength of a reason is positive, the reason is *prima facie* (using a term by Pollock); when it is positive, but smaller than 1, the reason is defeasible. For a defeasible reason  $(\varphi, \psi)$ , there exist circumstances  $\chi$  that defeat the reason, in the sense that  $(\varphi \wedge \chi, \psi)$  has strength zero. A *prima facie* reason  $(\varphi, \psi)$  can become weaker (‘diminished’ in Pollock’s terms; 2010) or stronger when the reason’s premises are extended to  $(\varphi \wedge \chi, \psi)$ . Note that a defeasible reason  $(\varphi, \psi)$  has an associated defeasible reason  $(\varphi, \neg\psi)$  for the opposite conclusion. As their strengths sum to 1, if one is weak, the other is strong, and vice versa.

<sup>6</sup>This notion of strength of a reason is not the same as the strength of a rule, as defined in Section 5.

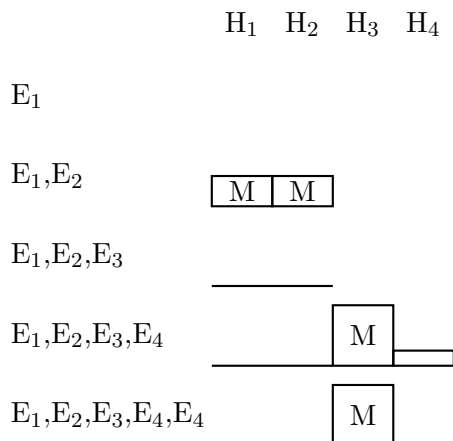


Figure 16: Development of the evidence, hypothetical scenarios and expectations

In many adversarial legal systems, the prosecution must prove its case ‘beyond a reasonable doubt’. In the present proposal, reasonable doubt can be thought of as the doubt that is made explicit in the model of the case. Reasonable doubt about a scenario exists as long as—according to the model about the case—the argument from all the evidence combined to the hypothetical scenario is not conclusive. There can also be doubt that is external to the model itself: perhaps the model is flawed, for instance when it was designed while ignoring alternative scenarios (tunnel vision), or perhaps the model needs to be reconsidered, for instance when newly found evidence sheds a different and unexpected light on the case. Since a good model includes all information that is considered to be relevant, doubt about a good model might be called ‘unreasonable’: there are no known reasons for the doubt.

## 6.2 A model of the running example according to this proposal

As an illustration of the proposal, we will develop a model analysing the investigation concerning the murder case used throughout this paper. Figure 16 illustrates the development of the evidence (on the left), the hypothetical scenarios (at the top) and the expectations (in the boxes). Each rectangle suggests a hypothetical scenario considered possible. In the model, at the end of the investigation, four scenarios have been considered: three murder scenarios  $H_1$ ,  $H_2$  and  $H_3$  (the first two about the asylum seekers, and the third about John), and one in which another male is the source of the profile found at the crime scene ( $H_4$ ). Some rectangles are ‘closed’—visually shown as a line—, as they represent scenarios that were initially considered possible but not anymore, as its probability dropped to 0 in the light of new evidence.

The first evidence considered, referred to as  $E_1$ , is the evidence based on the findings at the murder scene, such as the victim’s body, and the blood trace found on the victim’s coat. As yet very little is known about the scenario of how the crime developed, but one key expectation is already in place: by the finding of the blood trace found it is expected that the murderer’s DNA will match with that of the blood trace. Referring to the expected finding of the match as  $M$  and to a murder scenario  $H$  (i.e., in Figure 16 one of  $H_1$ ,  $H_2$  and  $H_3$ ), we therefore have:

$$P(M \mid H \wedge E_1) = 1. \tag{1}$$

The statement expresses that given the evidence  $E_1$  found at the murder scene, and assuming that the murder scenario  $H$  is true, the finding of a DNA match is expected as certain. In other words,  $H \wedge E_1$  is a conclusive reason for  $M$ , where  $M$  expresses the generic finding of a match, abstracting from the specific hypothetical scenario.

The murder scenarios considered become more specific when the two asylum seekers, one from Iraq, one from Afghanistan, become suspects in the investigation. We will write  $H_1$  and  $H_2$  to refer to the two scenarios in which one of these asylum seekers is Mary's murderer, and  $E_2$  for the evidence that led to the suspicion, perhaps not much more than the fact that they were seen in the neighbourhood. The two hypothetical scenarios are considered possible, which is indicated by their probability being positive given the evidence:

$$P(H_i | E_1 \wedge E_2) > 0, \text{ for } i = 1 \text{ and } i = 2. \quad (2)$$

In other words,  $E_1 \wedge E_2$  is a reason for  $H_1$  and for  $H_2$ . The strength of the reason is unknown, but positive.

Assuming that there is only one murderer, the two hypotheses are incompatible:

$$P(H_1 \wedge H_2 | E_1 \wedge E_2) = 0. \quad (3)$$

We can say that  $E_1 \wedge E_2$  excludes the conjunctive combination of the hypotheses  $H_1 \wedge H_2$ , and  $E_1 \wedge E_2$  is not a prima facie reason for  $H_1 \wedge H_2$ . It follows that the hypotheses exclude each other: for instance, it now holds that  $P(H_2 | H_1 \wedge E_1 \wedge E_2) = 0$ .

Since  $H_1$  and  $H_2$  are murder scenarios, we expect a DNA match:

$$P(M | H_i \wedge E_1 \wedge E_2) = 1, \text{ for } i = 1 \text{ and } i = 2. \quad (4)$$

It turns out that the DNA of the asylum seekers does not match with that of the blood sample ( $E_3$ ). So, provided that the conditional probability is defined, we have:

$$P(M | H_i \wedge E_1 \wedge E_2 \wedge E_3) = 0, \text{ for } i = 1 \text{ and } i = 2. \quad (5)$$

In other words,  $E_3$  provides defeating circumstances for the reason  $H_i \wedge E_1 \wedge E_2$  for  $M$ . Note that this shows that a reason that is conclusive can still be defeated, namely when adding information to its antecedent makes its consequent impossible.

Since we already had  $P(M | H_i \wedge E_1 \wedge E_2) = 1$  (for  $i = 1$  and  $i = 2$ ), it must follow that the conditional probability  $P(M | H_1 \wedge E_1 \wedge E_2 \wedge E_3)$  is undefined:

$$P(H_i \wedge E_1 \wedge E_2 \wedge E_3) = 0, \text{ for } i = 1 \text{ and } i = 2. \quad (6)$$

The two asylum seeker hypotheses are excluded, as  $H_1 \wedge E_1 \wedge E_2 \wedge E_3$  and  $H_2 \wedge E_1 \wedge E_2 \wedge E_3$  are not possible: neither of the two asylum seekers is the murderer.

Then the unexpected match with John's DNA is found during the extensive screening of the local population, making John the primary suspect. The hypothetical scenario in which John is the murderer is denoted  $H_3$ . In order to include the probabilistic information about a random match in our model, we need a fourth hypothetical scenario, namely that in which another male is the source of the profile ( $H_4$ ). For this scenario, we do not conclusively expect a match with John's DNA. This will only very rarely be the case, namely only with the probability of a random match:

$$P(M | H_4) = 0.66 \cdot 10^{-21}. \quad (7)$$

$H_4$  is a prima facie reason for finding a match with John's DNA, but a (very) weak one. A random person will only very rarely match John's DNA. It is hard to estimate what happens to this number when the evidence (preceding the finding of the match) is included, but it seems safe to assume that it remains positive:

$$P(M | H_4 \wedge E_1 \wedge E_2 \wedge E_3) > 0. \quad (8)$$

Using the terminology of the proposal,  $H_4 \wedge E_1 \wedge E_2 \wedge E_3$  is a prima facie reason for finding a match.

The match is established (evidence  $E_4$ ), hence  $P(M \mid H_4 \wedge E_1 \wedge E_2 \wedge E_3 \wedge E_4) = 1$ , even  $P(M \mid E_1 \wedge E_2 \wedge E_3 \wedge E_4) = 1$ .

It is now tempting—but fallacious—to conclude that  $H_3$  is (much) more probable than  $H_4$ , given the evidence. This could be expressed as follows:

$$P(H_3 \mid E_1 \wedge E_2 \wedge E_3 \wedge E_4) > P(H_4 \mid E_1 \wedge E_2 \wedge E_3 \wedge E_4). \quad (9)$$

However, this statement is not a formal consequence of the rest of the model, and would, if it is considered to be true, be a new assumption in the model.

Finally John confesses ( $E_5$ ), providing so many details of the crime and its circumstances that the confession is assessed as reliable. As a result of the confession, we add a new assumption to the model of the case, namely the certainty that John is the murderer:

$$P(H_3 \mid E_1 \wedge E_2 \wedge E_3 \wedge E_4 \wedge E_5) = 1. \quad (10)$$

$E_1 \wedge E_2 \wedge E_3 \wedge E_4 \wedge E_5$  is a conclusive reason for  $H_3$ , according to the model of the case. Note that this does not follow formally from the rest of the model of the case. In particular, we have not used Bayesian updating. At this stage, according to the model, no alternative scenario is possible anymore. Such a model can for instance be reasonable, when even the defence, a crucial source of relevant alternative scenarios to consider and to exclude, does not propose additional possibilities.

On the basis of this model of the running example, a final conclusion about who committed the crime can be drawn, since the combination of all the evidence provides a conclusive reason for the scenario according to which John committed the crime. Note that, in contrast with a Bayesian network analysis of the case, we have only specified some numbers, namely only those that we want to commit to: the qualitative numbers 1 and 0 corresponding to the logical truth values true and false, and the random match probability associated with the DNA analysis.

### 6.3 Strengths and limitations

Strengths and limitations of the proposal in this section are as follows:

#### Strengths

1. In the proposal, arguments for and against different scenarios, and the evidence expected given the truth of a scenario, are analysed within standard probability theory and its underlying classical logic.
2. Arguments can have different strengths, measuring degrees of uncertainty, with a strength of 1, representing conclusiveness. A scenario is considered as possibly true when its probability is modelled as positive, and excluded as a hypothesis when its probability is zero.
3. It is accepted that not all probabilities are available, or can reasonably be determined, by using a model of the case that partially specifies a probability function.

#### Limitations

1. In the present proposal, the arguments considered have an elementary structure, whereas contemporary approaches to defeasible argumentation develop an elaborate theory of argument structure.
2. No analysis is provided of argumentation schemes and scenario schemes.

## 7 Conclusion

In this paper, we have studied connections between arguments, scenarios and probabilities as normative frameworks in reasoning with evidence. Such a study is relevant given the different backgrounds of the people involved in criminal investigation and decision-making: Arguments and scenarios are familiar among fact-finders and lawyers, whereas probabilities are prominent in reports by forensic experts. By studying connections between arguments, scenarios and probabilities, we hope to contribute to the reduction of reasoning errors and miscommunication caused by these different backgrounds.

Our work builds on recent developments to study reasoning with forensic evidence probabilistically, and in particular using Bayesian networks (Taroni et al., 2006; Fenton, 2011). Since it is known that it is easy to misinterpret Bayesian networks, for instance causally (Dawid, 2010), we have started the exploration of the combined modelling of arguments and scenarios. Our approach continues earlier work on the design of structured probabilistic models and their explanation (Hepler et al., 2007; Fenton et al., 2013; Lacave and Díez, 2002; Levitt and Laskey, 2000; Druzdzel, 1996). Other research on modeling evidence using Bayesian networks is (Shen et al., 2006; Vreeswijk, 2005; Prakken and Renooij, 2001)

We reviewed research on the formal and computational connections between three normative frameworks for evidential reasoning based on arguments, scenarios and probabilities, respectively. In Sections 3 to 5, we studied pairwise connections, and in Section 6, connections between all three.

In Section 3, we discussed a hybrid model connecting arguments and scenarios. We saw how reasoning with arguments can be combined with reasoning to the best explanatory scenario, and the role of reasoning about evidential and causal rules. By the study of scenario schemes, we emphasised the relevance of the global coherence of scenarios. The modelling of degrees of uncertainty is not included in this proposal. The formalism is based on formal models of argumentation, of which the connections with standard formalisms, such as classical logic and standard probability theory, is as yet not fully understood.

In Section 4, the focus shifted to the combination of scenarios and probabilities. We showed how scenarios can be embedded in Bayesian networks, thereby connecting the role of the global coherence of scenarios with degrees of uncertainty. We showed how coherence, transfer of evidential support and the unfolding of scenarios can be modelled in Bayesian networks. In the approach, we encountered a standard limitation of Bayesian networks, namely that—since a Bayesian network is a model of a full probability function—more numbers are needed than are readily available or can reasonably be estimated. Also we noted that the resulting networks quickly become so large and complex that further efforts are needed to provide ways to explain such networks to professionals involved in criminal investigation and decision-making.

In Section 5, we studied connections between arguments and probabilities. We showed how arguments can be extracted from a Bayesian network. We formally defined the strengths of the rules and exceptions used to construct arguments. The arguments extracted from a Bayesian network can help explain such networks, even complex ones. The explanatory power is limited by the fact that our current algorithms produce arguments in many small variations. The argument construction algorithms are also computationally complex.

In Section 6, we proposed a view on arguments to and from scenarios in the context of probability theory. We studied arguments for and against different scenarios, using standard probability theory and its underlying classical logic. An argument can have a strength, measured by a conditional probability, which expresses a degree of uncertainty. It is not required that a full probability function is specified, as a model of a case only uses the numeric information that is available. The arguments studied in the proposal have an elementary structure, and no

analysis is provided of argumentation schemes and scenario schemes.

There are many remaining hard questions about the safe handling of probabilistic and non-probabilistic evidence in criminal investigation and decision-making. Still we hope that the lessons that we have learnt by studying the different connections between arguments, scenarios and probabilities, will gradually contribute to the prevention of reasoning errors, and a reduction of miscommunication between fact-finders and forensic experts.

## Acknowledgments

The research reported in this paper has been performed in the context of the project ‘Designing and Understanding Forensic Bayesian Networks with Arguments and Scenarios’, funded in the NWO Forensic Science program (<http://www.ai.rug.nl/~verheij/nwofs/>).

## References

- Anderson, T., Schum, D., and Twining, W. (2005). *Analysis of Evidence. 2nd Edition*. Cambridge University Press, Cambridge.
- Bench-Capon, T. J. M. and Dunne, P. E. (2007). Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641.
- Bennett, W. L. and Feldman, M. S. (1981). *Reconstructing Reality in the Courtroom*. London: Tavistock Feldman.
- Bex, F. J. (2011). *Arguments, Stories and Criminal Evidence: A Formal Hybrid Theory*. Springer, Berlin.
- Bex, F. J. (2014). Towards an integrated theory of causal scenarios and evidential arguments. In Parsons, S., Oren, N., Reed, C., and Cerutti, F., editors, *Computational Models of Argument. Proceedings of COMMA 2014*, pages 133–140. IOS Press, Amsterdam.
- Bex, F. J., Prakken, H., Reed, C. A., and Walton, D. N. (2003). Towards a formal account of reasoning about evidence: Argumentation schemes and generalisations. *Artificial Intelligence and Law*, 11(2/3):125–165.
- Bex, F. J., van Koppen, P. J., Prakken, H., and Verheij, B. (2010). A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law*, 18:1–30.
- Bex, F. J. and Verheij, B. (2013). Legal stories and the process of proof. *Artificial Intelligence and Law*, 21(3):253–278.
- Bondarenko, A., Dung, P. M., Kowalski, R. A., and Toni, F. (1997). An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101.
- Broeders, T. (2009). Decision making in the forensic arena. In Kaptein, H., Prakken, H., and Verheij, B., editors, *Legal Evidence and Proof: Statistics, Stories, Logic*, pages 71–92. Ashgate, Farnham.
- Buchanan, M. (2007). Conviction by numbers. *Nature*, 445(10):254–255.
- Console, L. and Theseider Dupré, D. (1994). Abductive reasoning with abstraction axioms. In Lakemeyer, G. and Nebel, B., editors, *Foundations of Knowledge Representation and Reasoning*, pages 98–112. Springer, Berlin.



- Crupi, V., Tentori, K., and Gonzales, M. (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science*, 74(2):229–252.
- Dawid, A. P. (2010). Beware of the DAG! In Guyon, I., Janzing, D., and Schölkopf, B., editors, *JMLR Workshop and Conference Proceedings: Volume 6. Causality: Objectives and Assessment (NIPS 2008 Workshop)*, pages 59–86. jmlr.org.
- Dawid, A. P., Twining, W., and Vasiliki, M., editors (2011). *Evidence, Inference and Enquiry*. Oxford University Press, Oxford.
- Derksen, T. and Meijsing, M. (2009). The fabrication of facts: The lure of the incredible coincidence. In Kaptein, H., Prakken, H., and Verheij, B., editors, *Legal Evidence and Proof: Statistics, Stories, Logic*, pages 39–70. Ashgate, Farnham.
- Druzdzel, M. J. (1996). Qualitative verbal explanations in Bayesian belief networks. *Artificial Intelligence and Simulation of Behaviour Quarterly*, 94:43–54.
- Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357.
- Evet, I., Jackson, G., Lambert, J. A., and McCrossan, S. (2000). The impact of the principles of evidence interpretation on the structure and content of statements. *Science and Justice*, 40(4):233–239.
- Fenton, N. E. (2011). Science and law: Improve statistics in court. *Nature*, 479:36–37.
- Fenton, N. E., Neil, M. D., and Lagnado, D. A. (2013). A general structure for legal arguments about evidence using Bayesian Networks. *Cognitive Science*, 37:61–102.
- Gordon, T. F., Prakken, H., and Walton, D. N. (2007). The Carneades model of argument and burden of proof. *Artificial Intelligence*, 171(10–15):875–896.
- Hájek, A. (2011). Interpretations of probability. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Stanford University.
- Hepler, A. B., Dawid, A. P., and Leucari, V. (2007). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability and Risk*, 6(1–4):275–293.
- Jensen, F. V. and Nielsen, T. D. (2007). *Bayesian Networks and Decision Graphs*. Springer, Berlin.
- Josephson, J. R. (2002). On the proof dynamics of inference to the best explanation. In MacCrimmon, M. and Tillers, P., editors, *The Dynamics of Judicial Proof. Computation, Logic, and Common Sense*, pages 287–305. Physica-Verlag, Heidelberg.
- Josephson, J. R. and Josephson, S. G. (1996). *Abductive Inference: Computation, Philosophy, Technology*. Cambridge University Press, Cambridge.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Penguin, London.
- Kaptein, H., Prakken, H., and Verheij, B., editors (2009). *Legal Evidence and Proof: Statistics, Stories, Logic (Applied Legal Philosophy Series)*. Ashgate, Farnham.
- Kirschner, P. A., Shum, S. J. B., and Carr, C. S. (2003). *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. Springer, Berlin.

- Lacave, C. and Díez, F. J. (2002). A review of explanation methods for Bayesian Networks. *Knowledge Engineering Review*, 17(2):107–127.
- Levitt, T. S. and Laskey, K. B. (2000). Computational inference for evidential reasoning in support of judicial proof. *Cardozo Law Review*, 22:1691–1731.
- Modgil, S. and Prakken, H. (2013). A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397.
- Pardo, M. S. and Allen, R. J. (2008). Juridical proof and the best explanation. *Law and Philosophy*, 27:223–268.
- Pearl, J. (1988). Embracing causality in default reasoning. *Artificial Intelligence*, 35:259–271.
- Pennington, N. and Hastie, R. (1993). Reasoning in explanation-based decision making. *Cognition*, 49(1–2):123–163.
- Pollock, J. L. (1987). Defeasible reasoning. *Cognitive Science*, 11(4):481–518.
- Pollock, J. L. (1995). *Cognitive Carpentry: A Blueprint for How to Build a Person*. The MIT Press, Cambridge (Massachusetts).
- Pollock, J. L. (2010). Defeasible reasoning and degrees of justification. *Argument and Computation*, 1(1):7–22.
- Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1(2):93–124.
- Prakken, H. and Renooij, S. (2001). Reconstructing causal reasoning about evidence: a case study. In Verheij, B., Lodder, A. R., Loui, R. P., and Muntjewerff, A. J., editors, *Legal Knowledge and Information Systems. JURIX 2001: The Fourteenth Annual Conference*, pages 131–142. IOS Press, Amsterdam.
- Rahwan, I. and Simari, G. R., editors (2009). *Argumentation in Artificial Intelligence*. Springer, Dordrecht.
- Schank, R. and Abelson, R. (1977). *Scripts, Plans, Goals and Understanding, An Inquiry into Human Knowledge Structures*. Lawrence Erlbaum, Hillsdale.
- Schneps, L. and Colmez, C. (2013). *Math on Trial: How Numbers Get Used and Abused in the Courtroom*. Basic Books, New York (New York).
- 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*, 5:87–117.
- Taroni, F., Aitken, C., Garbolino, P., and Biedermann, A. (2006). *Bayesian Networks and Probabilistic Inference in Forensic Science*. Wiley, Chichester.
- Taroni, F., Champod, C., and Margot, P. (1998). Forerunners of Bayesianism in early forensic science. *Jurimetrics*, 38:183–200.
- Thompson, P. (2013). Forensic DNA evidence. The myth of infallibility. In *Genetic Explanations: Sense and Nonsense*, pages 227–255. Harvard University Press.

- 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.
- Tillers, P. (2005). Picturing factual inference in legal settings. In Schuenemann, B., Tinnefeld, M. T., and Wittmann, R., editors, *Gerechtigkeitswissenschaft: Kolloquium aus Anlass des 70. Geburtstages von Lothar Philipps*. Berliner Wissenschafts-Verlag, Berlin.
- Timmer, S. T., Meyer, J. J., Prakken, H., Renooij, S., and Verheij, B. (2013). Inference and attack in Bayesian Networks. In Hindriks, K., de Weerd, M., van Riemsdijk, B., and Warnier, M., editors, *25th Benelux Conference on Artificial Intelligence (BNAIC 2013)*, pages 199–206. Delft University.
- Timmer, S. T., Meyer, J. J., Prakken, H., Renooij, S., and Verheij, B. (2014). Extracting legal arguments from forensic Bayesian Networks. In Hoekstra, R., editor, *Legal Knowledge and Information Systems: JURIX 2014: The Twenty-Seventh Annual Conference*, pages 71–80. IOS Press, Amsterdam.
- Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge University Press, Cambridge.
- van den Braak, S. W., van Oostendorp, H., Prakken, H., and Vreeswijk, G. A. W. (2008). Representing narrative and testimonial knowledge in sense-making software for crime analysis. In Francesconi, E., Sartor, G., and Tiscornia, D., editors, *Legal Knowledge and Information Systems: JURIX 2008: The Twenty-First Annual Conference*, pages 160–169. IOS Press, Amsterdam.
- van Eemeren, F. H., Garssen, B., Krabbe, E. C. W., Snoeck Henkemans, A. F., Verheij, B., and Wagemans, J. H. M. (2014a). Chapter 11: Argumentation in Artificial Intelligence. In *Handbook of Argumentation Theory*. Springer, Berlin.
- van Eemeren, F. H., Garssen, B., Krabbe, E. C. W., Snoeck Henkemans, A. F., Verheij, B., and Wagemans, J. H. M. (2014b). *Handbook of Argumentation Theory*. Springer, Berlin.
- Verheij, B. (2003a). Artificial argument assistants for defeasible argumentation. *Artificial Intelligence*, 150(1–2):291–324.
- Verheij, B. (2003b). DefLog: on the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation*, 13(3):319–346.
- Verheij, B. (2005). *Virtual Arguments. On the Design of Argument Assistants for Lawyers and Other Arguers*. T.M.C. Asser Press, The Hague.
- Verheij, B. (2009). The Toulmin argument model in artificial intelligence. or: How semi-formal, defeasible argumentation schemes creep into logic. In Rahwan, I. and Simari, G. R., editors, *Argumentation in Artificial Intelligence*, pages 219–238. Springer, Berlin.
- Verheij, B. (2014a). Arguments and their strength: Revisiting Pollock’s anti-probabilistic starting points. In Parsons, S., Oren, N., Reed, C., and Cerutti, F., editors, *Computational Models of Argument. Proceedings of COMMA 2014*, pages 433–444. IOS Press, Amsterdam.
- Verheij, B. (2014b). To catch a thief with and without numbers: Arguments, scenarios and probabilities in evidential reasoning. *Law, Probability and Risk*, 13:307–325.

- Vlek, C. S., Prakken, H., Renooij, S., and Verheij, B. (2013). Modeling crime scenarios in a Bayesian Network. In *The 14th International Conference on Artificial Intelligence and Law (ICAIL 2013). Proceedings of the Conference*, pages 150–159. ACM Press, New York (New York).
- Vlek, C. S., 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. S., Prakken, H., Renooij, S., and Verheij, B. (2014b). Extracting scenarios from a Bayesian Network as explanations for legal evidence. In Hoekstra, R., editor, *Legal Knowledge and Information Systems: JURIX 2014: The Twenty-Seventh Annual Conference*, pages 150–159. IOS Press, Amsterdam.
- Vreeswijk, G. A. W. (2005). Argumentation in Bayesian belief networks. In Rahwan, I., Moraitis, P., and Reed, C., editors, *Argumentation in Multi-Agent Systems*, volume 3366 of *Lecture Notes in Computer Science*, pages 111–129. Springer, Berlin.
- Wagenaar, W. A., van Koppen, P. J., and Crombag, H. F. M. (1993). *Anchored Narratives. The Psychology of Criminal Evidence*. Harvester Wheatsheaf, London.
- Walton, D. N., Reed, C., and Macagno, F. (2008). *Argumentation Schemes*. Cambridge University Press, Cambridge.
- Wigmore, J. H. (1913). *The Principles of Judicial Proof or the Process of Proof as Given by Logic, Psychology, and General Experience, and Illustrated in Judicial Trials. (Second edition 1931.)*. Little, Brown and Company, Boston (Massachusetts).