University of Groningen

A method for explaining Bayesian networks for legal evidence with scenarios

Vlek, Charlotte; Prakken, Hendrik; Renooij, Silja; Verheij, Bart

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

A method for explaining Bayesian networks for legal evidence with scenarios

Charlotte S. Vlek, Henry Prakken, Silja Renooij & Bart Verheij

ONLINE FIRST

Artificial Intelligence and Law

Vol. 24 No. 2
June 2016

ISSN 0924–8463

Springer

Springer

CrossMark

A method for explaining Bayesian networks for legal evidence with scenarios

Charlotte S. Vlek[1] · Henry Prakken[2,3] ·
Silja Renooij[2] · Bart Verheij[1]

Abstract In a criminal trial, a judge or jury needs to reason about what happened based on the available evidence, often including statistical evidence. While a probabilistic approach is suitable for analysing the statistical evidence, a judge or jury may be more inclined to use a narrative or argumentative approach when considering the case as a whole. In this paper we propose a combination of two approaches, combining Bayesian networks with scenarios. Whereas a Bayesian network is a popular tool for analysing parts of a case, constructing and understanding a network for an entire case is not straightforward. We propose an explanation method for understanding a Bayesian network in terms of scenarios. This method builds on a previously proposed construction method, which we slightly adapt with the use of scenario schemes for the purpose of explaining. The resulting structure is explained in terms of scenarios, scenario quality and evidential support. A probabilistic interpretation of scenario quality is provided using the concept of scenario schemes. Finally, the method is evaluated by means of a case study.

✉ Charlotte S. Vlek
c.s.vlek@rug.nl

Henry Prakken
h.prakken@uu.nl

Silja Renooij
s.renooij@uu.nl

Bart Verheij
b.verheij@ai.rug.nl

[1] Institute of Artificial Intelligence, University of Groningen, Groningen, Netherlands

[2] Department of Information and Computing Sciences, Utrecht University, Utrecht, Netherlands

[3] Faculty of Law, University of Groningen, Groningen, Netherlands

Springer

1 Introduction

In a criminal trial, a judge or jury draws a conclusion about what happened based on the evidence. This evidence often includes statistical information, especially with the increased use of forensic techniques such as DNA profiling. The task of a judge or jury is to take into account the whole case, including both statistical and non-statistical evidence. Whereas a probabilistic approach to reasoning with evidence is suitable for dealing with quantitative statistical information, in order to consider the perspective of the case as a whole, a judge or jury will be more inclined to use a narrative or argumentative perspective. In this paper, we propose to combine probabilistic and narrative techniques to allow a judge or jury to work with statistical information while considering the whole case in terms of narrative techniques.

Bayesian networks have become popular as a probabilistic tool for working with forensic evidence (see, e.g., Taroni et al. 2006). A Bayesian network consists of a graph and probability tables, together representing a joint probability distribution. From a Bayesian network, any probability of interest can be computed. The graphical structure contains information about (in)dependencies between variables, which makes a Bayesian network particularly suitable for modelling complex relations between variables. A Bayesian network is therefore potentially a good candidate for representing a case as a whole, since it deals well with interactions between variables. However, not much work has been done on modelling entire legal cases in a Bayesian network [a notable exception is Kadane and Schum (1996)]. Though a Bayesian network, once specified, can be very useful to perform computations with probabilistic information, the construction and understanding of a network are not straightforward. When formalizing the entire case in a probabilistic framework, it is crucial that the judge or jury gain some insight into the content of the model, lest they hand over the decision of the case entirely to the modeller.

A narrative approach to reasoning with legal evidence revolves around the concept of scenarios (or stories) (Bennett and Feldman 1981; Wagenaar et al. 1993; Pennington and Hastie 1993; Bex 2011). Forming scenarios about what may have happened is an intuitive way for jurors to make sense of the evidence, as was shown in experiments by Pennington and Hastie (1993). In a typical narrative approach, alternative scenarios are compared on key properties such as the quality of each scenario and how well each relates to the available evidence in terms of evidential support or evidential coverage. While in a probabilistic context the evidential support is fairly straightforward to interpret, the quality of a scenario is in general less well-defined.

In this paper, we aim to explain a Bayesian network in terms of scenarios. In particular, we aim to explain the content of the network by reporting which scenarios were modelled in a Bayesian network, their evidential support and their scenario quality. This way, the results of a network can be understood by a judge or jury in terms of narrative properties. This paper builds on previous work on the construction of Bayesian network graphs with scenarios (Vlek et al. 2014), and

extends previous preliminary ideas on explanation techniques using scenario schemes and scenario quality (Vlek et al. 2015a, b). In this paper we combine and further formalise these preliminary ideas, and use them to propose a reporting format for explaining a Bayesian network. Finally, we present a case study to test our proposed method.

The aim of the method proposed in this paper is to explain the content of a Bayesian network in terms of scenarios, scenario quality and evidential support. To this end, the previously proposed construction method is adapted with the use of scenario schemes, such that scenario quality can be represented in a Bayesian network. To evaluate our method, a case study is performed with the following criteria in mind:

1. Does the use of scenario schemes in the adapted construction method assist the modeller when constructing a Bayesian network?
2. Are various scenarios and their properties (e.g., quality) adequately represented in the Bayesian network?
3. Does the proposed explanation contain the required information about scenarios, scenario quality and evidential support?

To summarize, the contributions of this paper are as follows: (1) we present a method for explaining a Bayesian network with scenario schemes, (2) we present a probabilistic interpretation of scenario quality and (3) we evaluate the proposed method with a case study. The paper is organised as follows: background information on Bayesian networks, narrative and previous work is discussed in Sect. 2, the adaptation of the construction method with scenario schemes and scenario quality is discussed in Sect. 3, while explaining a network is discussed in Sect. 4. A case study is presented in Sect. 5, followed by a section on related work (Sect. 6) and a conclusion (Sect. 7).

2 Preliminaries

In this section we introduce Bayesian networks (Sect. 2.1), some previous work on the construction of Bayesian networks using a narrative approach (Sect. 2.2) and some concepts from scenario-based reasoning (the narrative approach, Sect. 2.3).
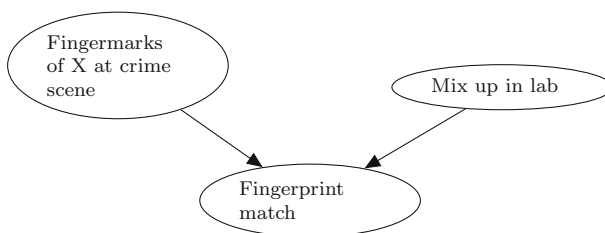
Fig. 1 An example of a Bayesian network graph

2.1 Bayesian networks

A Bayesian network represents a joint probability distribution (JPD) (Jensen and Nielsen 2007). It consists of a directed acyclic graph and probability tables for each node in the graph. An example of a Bayesian network graph is shown in Fig. 1. Each node in a Bayesian network represents a variable that can have several values such as, e.g., true/false (T/F), but more than two values are also possible. The probability table of a node V gives the conditional probabilities for that node taking each value given the values of its parents. For example, the probability table for `Fingerprint match` gives the probability $\Pr($`Fingerprint match|Fingermarks of X at crime scene, Mix up in lab`$)$ for all configurations of values of the three nodes. If a node V has no parents, the probability table specifies V's prior probability distribution.

The arrows in the graph can be modelled such that they represent causal relations, but they technically only represent a possible correlation (Dawid 2009). From the graph, information about (in)dependencies between variables can be read. Only the absence of an influence can be read from the network; the variables are then said to be d-separated. If two variables are not d-separated, they are said to be d-connected, and an influence is possible (but not necessarily the case). The concepts of d-connectedness and d-separation depend on observations that have been entered in the graph. For a serial connection $A \rightarrow C \rightarrow B$ or a diverging connection $A \leftarrow C \rightarrow B$ and no observations for node C, there is said to be an active chain between A and B, and the two nodes are d-connected. As soon as C is observed, the chain becomes inactive and when there are no other active chains between A and B, they are d-separated. The converging connection $A \rightarrow C \leftarrow B$ is an exception: the chain between A and B is blocked when C is not observed, yielding A and B d-separated when there are no other active chains between A and B. As soon as C or one of its descendants is observed, the chain becomes active and A and B are d-connected.

From a completely specified Bayesian network, any prior or posterior probability given the evidence can be calculated. In the example network from Fig. 1, one will typically enter an observation about the evidence (a match was found, `Fingerprint match = T`), and then calculate the probability of the hypothesis that the suspect X left a fingermark, given this evidence: $\Pr($`Fingermarks of X at crime scene = T|Fingerprint match = T`$)$.

2.2 Constructing a Bayesian network with idioms

The construction of a Bayesian network is not straightforward: firstly, the structure of the graph needs to be established, and secondly, the numbers in the probability tables need to be elicited. For methods to elicit the numbers, see, e.g., Renooij (2001). In this section we review methods for the construction of the graph, and in particular those that use idioms to assist a modeller with finding the structure of a Bayesian network.

In the legal domain, each case requires a custom Bayesian network. Nonetheless, some basic substructures recur throughout various cases, as pointed out by Hepler et al. (2004). By using such substructures as basic building blocks, the task of

constructing a Bayesian network graph is simplified. This idea was further extended to a list of legal idioms by Fenton et al. (2013). They developed a list of legal idioms, which included, for example, idioms for modelling the accuracy of evidence and for modelling an alibi. In addition to these legal idioms, we proposed to represent scenarios in a Bayesian network with the use of four narrative idioms (Vlek et al. 2014). These four narrative idioms are the scenario idiom, the subscenario idiom, the variation idiom and the merged scenarios idiom.

The main purpose of the scenario idiom, shown in Fig. 2, is to capture the coherence of a scenario in a Bayesian network: the fact that the elements of a scenario together form a coherent whole. Each element of the scenario is represented with a boolean node (E_i in Fig. 2). To represent the coherence of the scenario these nodes are clustered by adding a boolean parent node, called the scenario node, with arrows from the scenario node to each element of the scenario. Arrows are also drawn within the scenario (between elements of the scenario, represented as dashed lines in Fig. 2) to represent internal connections.

The scenario node represents the scenario as a whole and the arrows from the scenario node to each element of the scenario signify a special connection (drawn as double arrows in Fig. 2), namely that if the whole scenario is true, then each element must logically also be true. The numbers in the probability table of each element of the scenario are thus partially constrained such that the element is true with probability 1 if the scenario node is true, as shown in the table in Fig. 2. The coherence of the scenario is now captured with the scenario node, which guarantees that elements of the scenario are always d-connected since the scenario node itself is never observed. As a result of the constrained probabilities, when one element of the scenario becomes more probable, the other elements of the scenario become more probable as well due to an influence via the scenario node (in the absence of other influences). This is an effect of the coherence of a scenario, which in Vlek et al. (2014) is called the transfer of evidential support.

The subscenario idiom from Vlek et al. (2014) has a similar structure to the scenario idiom but is always used within a scenario. This idiom thus captures the coherence of a subscenario when it occurs within a larger scenario. The variation idiom is used to model small variations within a scenario, reducing the need for modelling many different scenarios that largely overlap. Finally, the merged scenarios idiom is used to combine all separate scenario structures as modelled with the scenario idioms, to produce one Bayesian network graph in which these

Fig. 2 The scenario idiom from Vlek et al. (2014). *Double arrows* signify the special relation between the scenario node and an element node, which results in the partial constraints for the probabilities of element nodes as shown in the table. *Dashed arrows* show possible connections

Scenario node	T	F
$E_1 =$T	1	...
$E_1 =$F	0	...
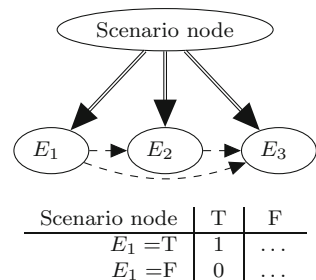
scenarios are mutually exclusive (but not necessarily exhaustive). Representing multiple scenarios in a single Bayesian network allows for a comparison between alternatives as is customary in the narrative approach, ensuring that probabilities are consistent with each other throughout the various alternative scenarios.

In addition to the four narrative idioms, in Vlek et al. (2014) we also introduced the concept of unfolding a scenario to gradually construct a Bayesian network. The method of unfolding relies on the idea that a scenario can have varying levels of detail, and more detail can be added when it is relevant for the specific case. A Bayesian network is then gradually constructed by starting with an initial scenario. More detail is added to this scenario by replacing elements of the scenario with subscenarios, where the subscenario node replaces the element node in the original scenario.

The four idioms and the method of unfolding led to the following design method in four steps from Vlek et al. (2014):

1. Collect: gather relevant scenarios for the case;
2. Unfold: for each scenario, model an initial scenario with the scenario idiom. Then unfold this scenario by repeatedly asking the three questions:

   a. Is there evidence that can be connected directly to the element node? If so, no unfolding is required.
   b. Is there relevant evidence for details of a subscenario for this element? If so, unfolding is required.
   c. Would it be possible to find relevant evidence for details of the subscenario for this element? If so, unfolding is required.

   Use the subscenario idiom to model the unfolding subscenarios and the variation idiom whenever a variation is encountered. The process of unfolding is finished when the three questions indicate that no more relevant evidence can be added to the structure;
3. Merge: use the merged scenarios idiom to merge the scenario structures constructed in the previous step;
4. Include evidence: for each piece of evidence that is available, include a node and connect it to the element node it supports. Additionally, include nodes for evidential data that is to be expected as an effect of elements in the structure.

2.3 Scenario-based reasoning

In a narrative approach to reasoning with legal evidence (see, e.g., Bennett and Feldman 1981; Wagenaar et al. 1993; Pennington and Hastie 1993; Bex 2011), several alternative scenarios are formulated and compared. To find which scenario is the best, the alternative scenarios are compared on two main aspects: the relations with the evidence (often called evidential coverage Pennington and Hastie 1993 or evidential support Bex 2011) and the quality of the scenario [also called coherence by Pennington and Hastie (1993) or goodness by Wagenaar et al. (1993)]. Whereas

the evidential coverage or evidential support is a fairly clear concept, the quality of a scenario is not as straightforward.

In their Anchored Narratives theory, Wagenaar et al. (1993) distinguish the need to anchor a scenario in common sense rules and to assess the goodness of the scenario. For this, they refer to Bennett and Feldman (1981), who say that a good scenario should have a clearly identifiable central action and a setting from which this central action seems plausible. According to Wagenaar et al. (1993), the properties of a good scenario are as follows: it should contain no contradictions, it should have a central action and other elements of the scenario should unambiguously explain how and why this central action took place.

According to Pennington and Hastie (1993), the quality of a scenario is split up into three main factors: completeness, consistency and plausibility. These three factors, which together form the scenario's coherence, should be assessed together with the evidential coverage and the uniqueness of the scenario. For a scenario to be consistent, it should contain no internal contradictions. Plausibility is the extent to which a scenario is in line with the decision maker's knowledge about the world. And a scenario is complete when it 'has all of its parts'. The latter relates to the idea that a good scenario has an underlying pattern, sometimes called a script (Schank and Abelson 1977) or a scenario scheme or story scheme (Bex 2011). A scenario scheme provides the abstract structure for a scenario, specifying which elements the scenario should contain. For Pennington and Hastie (1993), the abstract structure contains initiating states and events, goals, actions and consequences. More context-specific schemes are also used, such as by Schank and Abelson (1977), whose famous example is a restaurant script which specifies elements such as being seated, ordering, paying, et cetera.

In recent work, Bex (2011) formalized Pennington and Hastie's three factors (completeness, consistency, plausibility) in his hybrid theory of stories and arguments. In Bex's work, a scenario is complete when there is some scenario scheme such that for each element of the scheme, there is a corresponding element in the scenario (the scenario completes the scheme) and for each element in the scenario, there is some corresponding element in the scheme (the scenario fits the scheme). In Bex's work, a scenario is either consistent or inconsistent, and as soon as there is some internal contradiction in the scenario, it is inconsistent. Finally, the plausibility of a scenario depends on the plausibility of elements and connections within that scenario. This is formalized in terms of arguments: by forming arguments based on common sense knowledge, the elements and connections in a scenario can be supported or contradicted.

In this paper, we use the shared terminology of Pennington and Hastie, and Bex. We thus have the following three factors that determine the quality of a scenario:

- *Completeness* A scenario is complete when it 'has all its parts' (Pennington and Hastie 1993). This can be formalized with scenario schemes as done by Bex (2011). A scenario is then complete when there is some scenario scheme such that the scenario fits and completes the scheme.
- *Consistency* A scenario is consistent when it contains no internal contradictions.

– *Plausibility* The plausibility of a scenario is the extent to which it matches our common sense knowledge about the world. Bex (2011) speaks of the plausibility of elements and connections in a scenario, which is formalized using arguments. The plausibility of an element or connection is then the extent to which it is supported or attacked by arguments based on common sense knowledge.

3 Constructing a Bayesian network with scenario schemes and scenario quality

In this section, we use scenario schemes as a basis for context-specific idioms, which are called scenario scheme idioms. Each scenario scheme idiom provides a fixed structure in which only elements of the scenario need to be filled in, and they enable the explanation of the resulting network. The concept of a scenario scheme idiom builds on the scenario idiom as proposed in Vlek et al. (2014), but is now adapted based on preliminary work in Vlek et al. (2015a).

Three advantages of using scenario scheme idioms are as follows. Firstly, scenario scheme idioms provide more structure to the construction process. Secondly, the structure of these scenario scheme idioms is such that there is a close correspondence to the scenario itself. This allows for an explanation of the network, including the extraction of scenarios in text from the network, as will be discussed in Sect. 4.1. Thirdly, the use of scenario schemes makes it possible to represent the quality of a scenario in terms of completeness, consistency and plausibility as used by Pennington and Hastie (1993) and Bex (2011). This will be treated in Sect. 3.2 below, which formalises preliminary work from Vlek et al. (2015b). In Sect. 3.3, an updated design method is presented, which includes the use of scenario schemes and scenario quality to construct a Bayesian network.

3.1 Scenario scheme idioms

A scenario scheme provides the abstract structure of a scenario. For example, one can think of a typical murder scheme, which includes a motive, opportunity, killing and death of the victim. With a scenario scheme that lays out the structure, a scenario can be formed by filling in the specific propositions of which the scenario consists. With the example murder scheme, a scenario can be formed by specifying exactly what the motive is (e.g., the suspect wanted money), how the opportunity was created (e.g., the suspect was at the victim's home), et cetera. In what follows, we assume that such scenario schemes exist for various crime scenarios.

In a Bayesian network, the concept of scenario schemes can be used as building blocks for a Bayesian network. To this end, we use scenario scheme idioms, which are defined as follows:

Definition 1 (*Scenario scheme idiom*) A *scenario scheme idiom* is a specific Bayesian network fragment of which the graph $(\mathcal{V}, \mathcal{E})$ and the probabilities are constrained as follows:

- $\mathcal{V}$ consists of a boolean scenario node SCN which represents the scenario as a whole and boolean nodes V_i which each represent an element of the scenario scheme; and
- $\mathcal{E}$ consists of unlabelled connections (SCN, V_i) from SCN to each element node V_i (drawn as double arrows); and
- $\mathcal{E}$ possibly contains labelled connections (V_i, V_j, x) between element nodes V_i and V_j with label x. The label $x \in \{c, t\}$ indicates whether a connection is causal (c) or temporal (t); and
- The probability table for each element node V_i is constrained such that for any assignment to parent nodes $\text{pa}_S(V_i)$ of V_i within the same scenario,

$$\Pr(V_i = \text{T} | \text{SCN} = \text{T}, \text{pa}_S(V_i)) = 1.$$

As an example, a murder scheme idiom is shown in Fig. 3. Similar to the scenario idiom from Vlek et al. (2014) (see also Sect. 2), the scenario scheme idiom specifies the graphical structure by which a scenario can be represented in a Bayesian network, and it constrains some probabilities but not all. In particular, the double arrows from the scenario node (SCN) to each element node indicates the special connection between the scenario as a whole and its elements. The probability tables for element nodes are constrained such that given that the scenario node is true (SCN = T), the element node must be true with probability 1. This leaves the probabilities conditioned on SCN = F, indicating the probability of an element to be true when the scenario as a whole is not true (conditioned on any other parents that node might have), to be elicited by the modeller.

Whereas the previously developed scenario idiom provided a general clustered structure for representing any type of scenario, a scenario scheme idiom gives a specific structure for a scenario about a certain topic, such as a murder or a burglary. Elements of the scenario scheme idiom are used as 'place holders' for specific states
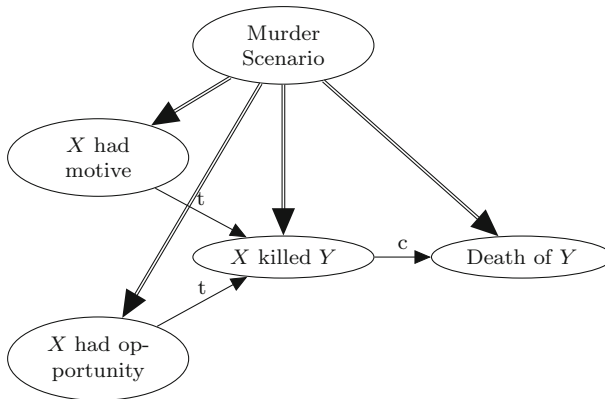
Fig. 3 The murder scheme idiom. *Double arrows* signify that the underlying probabilities are constrained

and events in a scenario. Moreover, while the previously developed scenario idiom had dashed arrows between element nodes to indicate possible connections, the scenario scheme idiom has fixed connections since it is used to represent a scenario with a specific structure. These connections are annotated with labels 'c' and 't', which provide more information for explaining the network at a later stage. It is of course possible to adapt these definitions such that scenarios can include more labels than just these two.

By specifying the elements of the scenario scheme idiom (e.g., the suspect wanted money for the motive), a scenario can be modelled in the network. The propositions of a scenario are then represented as boolean variables V_i with values true (T) and false (F), such that the value-assignment $V_i = T$ (e.g. `The suspect wanted money = T`) corresponds to the proposition in the scenario (the suspect wanted money).

When scenarios are modelled with a scenario scheme idiom, the network has the following properties by construction:

1. Boolean element nodes V_i relate to propositions in a scenario such that the proposition corresponds to the node taking value true $V_i = T$;
2. Arrows are annotated to show whether a connection is causal ('c') or temporal ('t');
3. Arrows correspond to the connections within the scenario.

With these properties, it becomes easier to understand the resulting network, since the propositions that make up a scenario and the connections between them are represented in an intuitive way. In Sect. 3.3 we present a design method in which scenario scheme idioms are used to construct a Bayesian network.

3.2 Representing scenario quality

The concept of scenario scheme idioms from the previous subsection facilitates the representation of a scenario in a Bayesian network. It also provides us with the tools to represent the quality of a scenario in the Bayesian network, in terms of completeness, consistency and plausibility (see Sect. 2.3).

3.2.1 Completeness

The completeness of a scenario was formalized by Bex (2011) using scenario schemes: a scenario is complete when it fits and completes a scenario scheme (see also Sect. 2.3). Inspired by Bex's definitions, we define the concept of completeness using scenario scheme idioms as follows:

Definition 2 (*Fitting a scenario scheme idiom*) A scenario *fits* a scenario scheme idiom when for every proposition p in the scenario, there is a corresponding node V in the scenario scheme idiom.

Definition 3 (*Completing a scenario scheme idiom*) A scenario *completes* a scenario scheme idiom when for every node *V* in the scenario scheme idiom there is some corresponding proposition *p* in the scenario.

This leads to the following definition of completeness, related to scenario scheme idioms:

Definition 4 (*Completeness*) A scenario is *complete* when it fits and completes a scenario scheme idiom.

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

As an example of an incomplete scenario that might be encountered during the modelling process, consider the following: 'Jane stabbed Mark, Mark died'. This scenario could fit the scenario scheme from Fig. 4a, but it lacks elements about a fight and about Jane having a knife. A modeller can complete this scenario by adding these elements as shown in Fig. 4b.

Note that the same incomplete scenario could be completed differently when a different scenario scheme idiom is used. For example, the incomplete scenario could also fit a scenario scheme idiom in which Jane stabbing Mark was self-defence. By adding the elements to complete the scenario relative to this scenario scheme idiom, the modeller can model this as a second scenario. Note that the modeller can model as many completed scenarios as are deemed relevant for the case; they are simply modelled as alternative possibilities and the model will be used to then find which scenario is most probable.
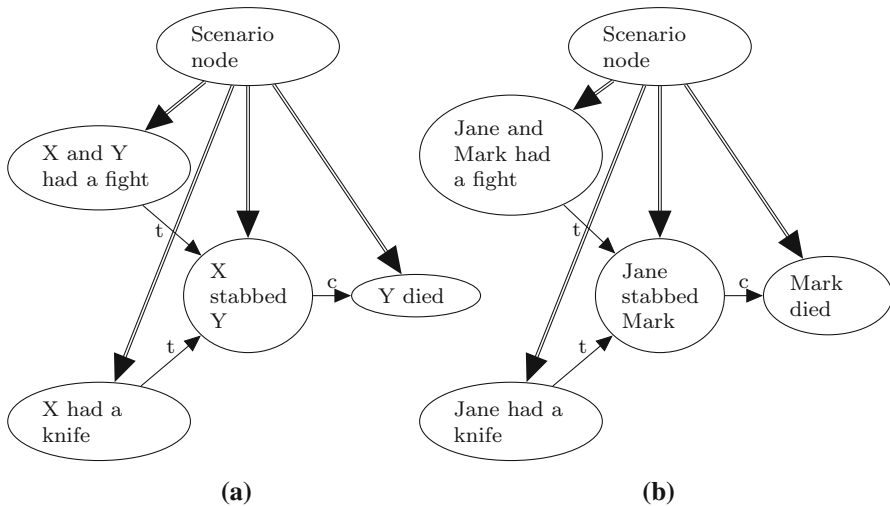
Fig. 4 A scenario scheme idiom and a corresponding network structure. **a** A scenario scheme idiom for a fight resulting in stabbing. **b** A network structure for the scenario about Jane stabbing Mark

3.2.2 Consistency

When a scenario is *inconsistent*, this is because two or more elements of the scenario are together inconsistent (see also Sect. 2.3). This can be represented in the network by adding a constraint node construction (see, e.g., Fenton et al. 2011) to the nodes that are together inconsistent. This way, the probabilities are set such that the elements that are together inconsistent ($V_1, V_2 \ldots$) have a probability of 0 to be true at the same time ($\Pr(V_1 = \text{T}, V_2 = \text{T}, \ldots) = 0$). Consistency and inconsistency can thus be defined as follows:

Definition 5 (*Inconsistency*) A scenario is *inconsistent* when there is a set of nodes $\{V_1, \ldots V_n\}$, with $n \geq 2$, corresponding to propositions in the scenario for which $\Pr(V_1 = \text{T}, \ldots, V_n = \text{T}) = 0$.

Definition 6 (*Consistency*) A scenario is *consistent* when it is not inconsistent.

As an example of an inconsistency in a scenario, consider a scenario that contains the elements 'Jane was at the cinema at 8 p.m.' and 'Jane was at Mark's place at 8 p.m.'. To represent this inconsistency, a constraint node is added to the network, and arrows are drawn from 'Jane was at the cinema at 8 p.m.' and 'Jane was at Mark's place at 8 p.m.' to the constraint node. The probability table of the constraint node is such that the constraint node has value `NA` (not applicable) whenever 'Jane was at the cinema at 8 p.m.' and 'Jane was at Mark's place at 8 p.m.' both have value `true`. By adding the evidence that the constraint node can never take value `NA`, the situation in which the two inconsistent elements are both true can never occur.

When a scenario's inconsistency is modelled with a constraint, it results in a probability of 0 for the scenario node of that scenario. To see this, consider a scenario with two inconsistent elements V_1 and V_2. If the scenario node had value `true`, these elements V_1 and V_2 would also have value `true` with probability 1. Since the constraint forces that they cannot both be true, it follows that the scenario node must have value `false`. A scenario which contains an inconsistency is thus not considered a viable alternative.

3.2.3 Plausibility

The plausibility of a scenario depends on the plausibility of the elements and connections in that scenario. What matters is how well an element or connection matches our common sense knowledge of the world (see Sect. 2.3 for more details). In our Bayesian network context, the common sense knowledge is arguably in the probabilities of the network: if our degree of belief in an element is high, this element is close to our common sense knowledge of the world. When the degree of belief in an element is low, we find it less credible based on our common sense knowledge. This leads to the following definitions of the plausibility of elements and connections:

Definition 7 (*Plausibility of an element*) The *plausibility of an element p* in a scenario, corresponding to node V, is given by $\Pr(V = \text{T})$.

Definition 8 (*Plausibility of a connection*) The *plausibility of a connection* between elements p_1 and p_2 in a scenario, with corresponding nodes V_1 and V_2, is given by $\Pr(V_2 = \text{T} | V_1 = \text{T})$.

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

To represent the plausibility of an element or connection in a Bayesian network, the modeller can first establish how plausible he or she finds a certain element of a scenario and then calculate the required numbers for the probability tables based on this. As an example, consider the implausible event 'Mark fell on a knife'. To reflect this implausibility, the probability $\Pr(\texttt{Mark fell on a knife} = \text{T})$ should be a low value. That probability will be calculated from the probability tables as follows (with the scenario node abbreviated to $\texttt{ScN}$ and in the absence of other parents):

$$\Pr(\texttt{Mark fell on a knife} = \text{T})$$
$$= \Pr(\texttt{Mark fell on a knife} = \text{T} | \texttt{ScN} = \text{T}) \cdot \Pr(\texttt{ScN} = \text{T})$$
$$+ \Pr(\texttt{Mark fell on a knife} = \text{T} | \texttt{ScN} = \text{F}) \cdot \Pr(\texttt{ScN} = \text{F}).$$

Since $\Pr(\texttt{Mark fell on a knife} = \text{T} | \texttt{ScN} = \text{T}) = 1$ by definition of the scenario scheme idiom, the probability $\Pr(\texttt{Mark fell on a knife} = \text{T} | \texttt{ScN} = \text{F})$ can be calculated based on $\Pr(\texttt{Mark fell on a knife} = \text{T})$ as soon as the probability table of the scenario node has been elicited with the probabilities $\Pr(\texttt{ScN} = \text{T})$ and $\Pr(\texttt{ScN} = \text{F})$. $\Pr(\texttt{Mark fell on a knife} = \text{T} | \texttt{ScN} = \text{F})$ is then found as follows:

$$\Pr(\texttt{Mark fell on a knife} = \text{T} | \texttt{ScN} = \text{F})$$
$$= \frac{\Pr(\texttt{Mark fell on a knife} = \text{T}) - \Pr(\texttt{ScN} = \text{T})}{\Pr(\texttt{ScN} = \text{F})}.$$

When a node has multiple parents, rather than speaking of the plausibility of a connection, one can speak of the plausibility of that element given its parents. Again, a modeller can establish the plausibility and then calculate the numbers for the probability table of a node with parents. Consider for example the connection between events 'Mark fell on a knife' and 'Mark died', which is considered implausible because it would be very unlikely that the knife was positioned in such a way that falling on it would result in a fatal injury. The element 'Mark died' is thus implausible given element 'Mark fell on a knife'. An appropriate probability needs to be set for $\Pr(\texttt{Mark died} = \text{T} | \texttt{Mark fell on a knife} = \text{T}, \texttt{Scenario node} = \text{F})$ in the probability table to represent this plausibility. Again, this can be calculated from the desired value for $\Pr(\texttt{Mark died} = \text{T} | \texttt{Mark fell on a knife} = \text{T})$, assuming that the probabilities for $\texttt{Mark fell on a knife}$ are already known. The calculation is then done as follows (with $\texttt{Mark fell on a knife}$ abbreviated to $\texttt{Mark fell...}$):

$$\text{Pr}(\text{Mark died} = \text{T}|\text{Mark fell on a knife} = \text{T}, \text{ScN} = \text{F})$$

$$= \frac{\text{Pr}(\text{Mark died} = \text{T}|\text{Mark fell...} = \text{T}) \cdot \text{Pr}(\text{Mark fell...} = \text{T}) - \text{Pr}(\text{ScN} = \text{T})}{\text{Pr}(\text{Markfell...} = \text{T}|\text{ScN} = \text{F}) \cdot \text{Pr}(\text{ScN} = \text{F})}$$

In sum, the modeller can go through the network starting at the root nodes to fill in the probability tables to represent plausibility. For a node V_1 with no parents within the scenario, the plausibility $\text{Pr}(V_1 = \text{T})$ can be used to calculate $\text{Pr}(V_1 = \text{T}|\text{ScN} = \text{F})$, as soon as the probabilities of ScN have been specified. Similarly, after specifying the probability table of V_1, a node V_2 connected to V_1 can be specified using the plausibility of V_2 given V_1. This way, the concept of plausibility is represented in the network.

3.3 Design method

With the adapted construction method as discussed in the previous subsections, a Bayesian network can be constructed such that it represents the available scenarios and their quality. In Vlek et al. (2014), a design method was presented in four steps, which can now be enhanced to also include the use of scenario schemes and scenario quality. This leads to two new steps: step 2, in which scenario scheme idioms are used to represent scenarios in a network, taking into account completeness and consistency, and step 6 in which numbers are specified, taking into account plausibility. The full design method is then as follows:

1. *Collect:* gather relevant scenarios for the case;
2. *Represent:* for each scenario, find one or more scenario scheme idioms to which the scenario fits. If the scenario is incomplete with respect to a scenario scheme idiom, extend the scenario such that it completes the scheme. Use the scenario scheme idioms to model each scenario. If the scenario is inconsistent, add a constraint node to model the inconsistency;
3. *Unfold:* for each scenario, unfold to more detailed subscenarios if needed by repeatedly asking the three questions:

   (a) Is there evidence that can be connected directly to the element node? If so, no unfolding is required.
   (b) Is there relevant evidence for details of a subscenario for this element? If so, unfolding is required.
   (c) Would it be possible to find relevant evidence for details of the subscenario for this element? If so, unfolding is required.

   Use the subscenario idiom to model the unfolding subscenarios and the variation idiom whenever a variation is encountered. The process of unfolding is finished when the three questions indicate that no more relevant evidence can be added to the structure;

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

5. *Include evidence:* for each piece of evidence that is available, include a node and connect it to the element node it supports. Additionally, include nodes for evidential data that is to be expected as an effect of elements in the structure.
6. *Specify numbers:* represent plausibility by calculating the corresponding probabilities for elements in the scenario. Specify other probabilities using elicitation techniques.

4 Explaining Bayesian networks with scenarios

When modelling an entire case, it is crucial that the judge or jury gains some understanding of the network, since the way in which a case is modelled determines the outcome of the network. It is therefore especially important that a judge or jury understands which assumptions go into a model (Fenton and Neil 2011), and in our approach this means: understanding which scenarios have been modelled and how they have been modelled when it comes to narrative properties such as evidential support and scenario quality. In particular, we aim to highlight the information (e.g., about scenario quality) that can help a judge or jury to understand the resulting calculations from the network. In this section, we discuss the ingredients needed for such a report: a method for extracting scenarios from a Bayesian network [Sect. 4.1, based on ideas from Vlek et al. (2015a)], how to calculate evidential support from the network [Sect. 4.2, also discussed in Vlek et al. (2015a)], and how to report about the quality of a scenario (Sect. 4.3). Finally, we use these ideas to propose a new reporting format in Sect. 4.4.

4.1 Extracting scenarios

Due to the construction with scenario scheme idioms, the resulting network is structured such that the scenarios represented in it can be identified easily. In particular, the scenarios in the network satisfy the following properties:

1. Each scenario is modelled as a cluster of nodes connected to a scenario node. The elements of a scenario are represented in the network by the child nodes of the corresponding scenario node (except the constraint node, which is also a child node of the scenario node). Similarly, elements of a subscenario are children of the subscenario node, et cetera.
2. The elements of a scenario (propositions) are modelled in the network as value assignments $V_1 = \text{T}$. For simplicity, we can simply write the name of the node V_1 as a proposition (omitting the value, which is implied).
3. Arrows between elements of the scenario represent connections within the scenario, and are labelled 't' (temporal) and 'c' (causal) to annotate additional information.

For a given network, the scenario in text form can be extracted by identifying which nodes represent elements of that scenario (using property 1), writing down the

propositions corresponding to these nodes (using property 2) and adding connections (with connectives 'then' and 'therefore', using property 3, as will be explained below). However, the propositions need to be ordered somehow, depending on the structure within the scenario. In Sect. 4.1.1, some examples are given of structures that may be encountered, and it is discussed how they can be ordered in the text. In Sect. 4.1.2, we propose a method for systematically producing such an ordering on the propositions for any structure, and we show how this method produces the desired results for the examples in Sect. 4.1.1.

4.1.1 Examples of scenario structures

We propose to translate a Bayesian network structure to text such that the directions of arrows are respected. This means that for nodes X and Y with an arrow $X \rightarrow Y$ (labelled c or t), the proposition corresponding to X precedes that corresponding to Y in the text. Note that such a linear ordering respecting the arrows this way is always possible, since a Bayesian network is acyclic.

Some examples of typical structures that may occur are shown in Fig. 5: the structure of a scenario can be linear (5a) or a node can have multiple parents (or children) (5b). Furthermore, paths of parents and children can be of varying lengths (5c) and (5d) or there can be connections between parents (or children) (5e).

A linear structure such as the one shown in 5a can be translated to 'A. Then B. Therefore C.'. The structure in 5b is an example of a node with two parents. In order to respect both arrows, parents are grouped together (with 'and' or 'or', see below), to produce, for instance, "A and B. Therefore C." Note that while the arrows have different labels, we used 'therefore' as a connecting word, since we think the causal connection is important to convey to the reader. We propose that whenever parents of one node are grouped and at least one is connected to the child node with label 'c', the connection is translated to 'therefore'.

Whether multiple parents are grouped with 'and' or 'or' depends on the probabilities. We propose to use 'and' as a default, but one situation in which 'or' is more appropriate, is when the so-called 'explaining away' effect occurs (Wellman and Henrion 1993). This is when the two parents A and B (now interpreted as causes of child C) form alternative explanations of C. Consider for example A: Jane had a knife and C: Jane stabbed Mark, but now with the addition of an alternative cause B:
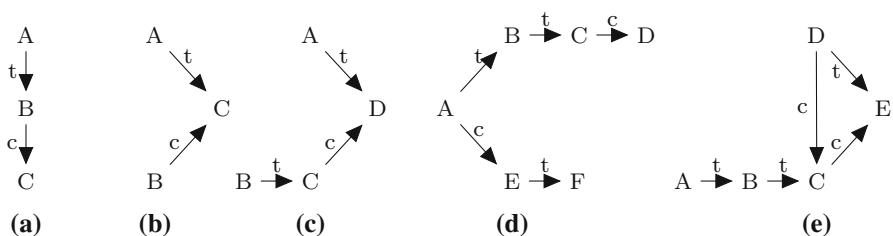
Fig. 5 Some typical structures

a knife was lying nearby. In this case, knowing A (Jane had a knife) reduces the need to assume B (a knife was lying nearby) to explain the effect C (Jane stabbed Mark). This effect of explaining away occurs when the effect on C of knowing B to be true is smaller when it is already known that A is true,, that is, when the following holds (Wellman and Henrion 1993):

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

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

Similar to the situation with multiple parents, we propose to group multiple children using 'and'. In this case, no distinction needs to be made between 'and' and 'or', because there is no influence between children unless there is a direct connection between them, which leads to a different situation (see below).

In Fig. 5c, d, there are multiple parents (c) or children (d), but with paths of varying lengths. Alternative translations for Fig. 5c are:

1. "B then C, and A. Therefore D."
2. "B. Then C and A. Therefore D."

The first translation can be ambiguous, and might lead to complicated combinations for longer paths (e.g., "K then L therefore M then N, and O. Therefore P"). We thus propose to use the second translation, which deals better with a structure as shown in Fig. 5d. There, the translation would be: "A. Therefore B and E. Then C and F. Therefore D." In general, this means that we group nodes together that are the same distance from a common ancestor or descendant. Note that this means that different structures may produce the same text: the translation of Fig. 5d would be the same if D were a child of F instead of C.

Finally, a situation is possible in which multiple parents or children also have a connection between them, as shown in Fig. 5e. There, E has parents C and D, while D is also a parent of C. In order to respect the arrows, D needs to precede C. We propose the following translation: "D. Therefore C. Therefore E". In this translation, the arrow from D to E is respected, but is not explicitly mentioned in the text. Including the remainder of the graph, the translation becomes: "A. Then B and D. Therefore C. Therefore E." Nodes B and D are grouped together because they are at the same distance from common descendant C.

The structures in Fig. 5 are examples of structures that may occur within a scenario. A special situation occurs when an element in a scenario is also the subscenario node for a subscenario. An example is shown in Fig. 6: there is now a subscenario for 'Mark died'. When such a subscenario occurs, we propose to maintain the concept of 'unfolding' within the text: in the main scenario the subscenario is represented as a single element (Mark died), but upon request it can be unfolded to a subscenario. For now, we represent the two versions of an element/subscenario by putting it in parentheses, in which the element (Mark died) is followed by the subscenario that can be shown upon request, as follows: [Mark died: Mark lost a lot of blood. Therefore, Mark died of blood loss]). In a user

interface the presentation should be such that only the element (Mark died) is shown, with some possibility to unfold to the subscenario.

4.1.2 A method for extracting text from any structure

In this section we propose a method for producing a text from any scenario structure. This method produces the desired translations as discussed in Sect. 4.1.1. Most importantly, the direction of arrows in the network are respected, such that when $X \rightarrow Y$, then X precedes Y in the text. Additionally, multiple parents or children of a node are grouped together when possible.

The method considers only the elements nodes in a scenario, excluding the scenario node. The idea is to start with the root nodes in this element subgraph (these root nodes are the elements of the scenario that have no parents *within the scenario*), and trace through connections to form sequences of elements. When a common descendant is found, the sequences are integrated to form one sequence such that nodes at the same distance of that common descendant are grouped together. Similarly, when a path splits, children are grouped together. The method is as follows:

1. Start with all nodes that have no parents within the scenario.
2. For a starting node X, find all children and the label ('c' or 't') of their connection to X. Write down a sequence starting with 'X.' followed by 'Then' when all children are connected with a 't'-connection and 'Therefore' when there is at least one 'c'-connection. When there are multiple children of X, add these to the sequence conjoined with 'and'. For a parent X with children Y and
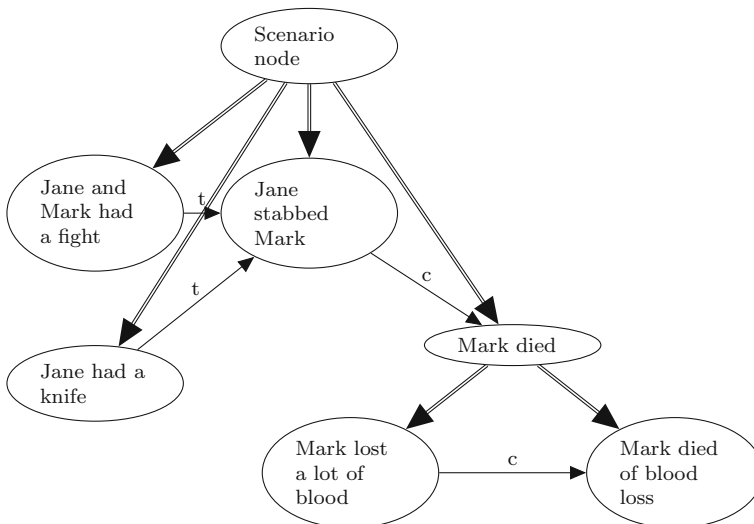
Fig. 6 A scenario with a subscenario

Z (where at least one connection is labelled 'c'), this becomes: "X. Therefore Y and Z."

3. For the set of children of the previous step, find all children of these nodes and conjoin them with 'and'. Place this at the end of the sequence, preceded by 'Therefore' if there is at least one 'c'-connection, and 'Then' otherwise.

4. As soon as a node is encountered that also occurs elsewhere in the sequence or in another sequence for that network, this is a common descendant. Integrate sequences pairwise as follows, repeating to form one sequence:

   (a) Each sequence consists of groups of nodes (possibly consisting of one node, or several nodes combined with 'and' or 'or'), separated by connectives 'therefore' or 'then'. Count the length of the sequences by counting the number of groups. Suppose sequence 1 is of length n and sequence 2 is of length k. Without loss of generalization, assume $n \geq k$. Then $d = n - k$ is the difference in length.

   (b) Merge the sequences by starting with the first d nodes of sequence 1, and whatever connectives are between them.

   (c) After that, group the $d + 1$st group of sequence 1 with the first group of sequence 2 (conjoined with 'and' or 'or', depending on the probabilities, see Sect. 4.1.1), the $d + 2$nd group of sequence 1 with the second group of sequence 2, et cetera, until the $n - 1$st group of sequence 1 is grouped with the $k - 1$st group of sequence 2.

   (d) Connectives in the new sequence are of type c whenever there is at least one c-connection between elements in one of the sequences, and otherwise t.

   (e) The last node of each sequence will be the common descendant, place this node in the sequence only once.

5. Repeat the process by expanding each sequence by finding child nodes and integrate whenever a common descendant is found.

This procedure gives the required results for the examples from Sect. 4.1.1. To see this, consider the structures from Fig. 5 once again (where 'and' is used as a default for multiple parents, since probabilities are not considered in this example):

5(a) The start node is A. Since the structure is linear, the following sequence is produced: 'A. Then B. Therefore C.'

5(b) There are two start nodes: A and B. This leads to two initial sequences 'A. Then C.' and 'B. Therefore C'. Node C is in both sequences, so they are integrated to obtain 'A and B. Therefore C'.

5(c) Again, A and B are the two start nodes. Two sequences 'A. Then D.' and 'B. Then C. Therefore D' are formed. These are of different lengths, so the final sequence becomes: 'B. Then A and C. Therefore D.'

5(d) There is one start node, A. Children of A are grouped together, and children of the set of children of A are again grouped together. This leads to 'A. Therefore B and E. Then C and F. Therefore D.'

5(e)   In this structure, A and D are both start nodes. Two sequences are formed to find one common descendant C: 'A. Then B. Then C.' and 'D. Therefore C'. These are integrated to obtain 'A. Then B and D. Therefore C.' Proceeding with this sequence, the last node E is added to obtain: 'A. Then B and D. Therefore C. Therefore E.'

Applied to the example from Fig. 6, this method proceeds as follows:

– Start with root nodes 'Jane and Mark had a fight' and 'Jane had a knife'.
– This results in two sequences with a shared node 'Jane stabbed Mark:

Jane and Mark had a fight. Then Jane stabbed Mark.
Jane had a knife. Then Jane stabbed Mark.

The two sequences are integrated:

Jane and Mark had a fight and Jane had a knife. Then Jane stabbed Mark.

– Another child is added to this sequence:

Jane and Mark had a fight and Jane had a knife. Then Jane stabbed Mark. Therefore, Mark died.

A similar process is done for the subscenario 'Mark died', which ultimately leads to the following text:

Jane and Mark had a fight and Jane had a knife. Then Jane stabbed Mark. Therefore, [Mark died: Mark lost a lot of blood. Therefore, Mark died of blood loss].

4.2 Evidential support

A probabilistic approach lends itself very well to analysing evidential support, which is often reported in terms of likelihood ratios (LR) (Taroni et al. 2006). However, a likelihood ratio is most useful when two mutually exclusive and jointly exhaustive hypotheses are compared (Fenton et al. 2014). This is not always the case when working with scenarios. Moreover, to adhere to the original narrative approach to reasoning with legal evidence, we aim to report the evidential support for each scenario separately, without comparing the scenarios to each other just yet (since for a comparison, scenario quality should also be taken into account).

Our goal is thus to report for each scenario how the evidence supports or attacks that scenario. For this, what matters is how the probability of a scenario with scenario node ScN changes after evidence e is added. So we want to compare the posterior probability $\Pr(\text{ScN} = \text{T}|e)$ with respect to the prior probability $\Pr(\text{ScN} = \text{T})$. This is expressed in the following measure for evidential support:

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

Several other measures for evidential support exist which also compare how the posterior probability changes with respect to the prior probability. Note that $\frac{\Pr(e|\text{ScN}=\text{T})}{\Pr(e)}$, is equal to our measure, as follows from Bayes' theorem. With our proposed measure, we have the following definition of evidential support/attack:

Definition 9 (*Evidential support*) For a scenario with scenario node ScN and evidence e, the *evidential support* of e for ScN is given by

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

Evidence e is

- supporting evidence if

$$\frac{\Pr(\text{ScN} = \text{T}|e)}{\Pr(\text{ScN} = \text{T})} > 1$$

- attacking evidence if

$$\frac{\Pr(\text{ScN} = \text{T}|e)}{\Pr(\text{ScN} = \text{T})} < 1$$

- neutral evidence if

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

In this definition, e can be one piece of evidence or a set of evidence. Reporting something about a set of evidence is important as well, because the evidential support cannot be expected to simply multiply for multiple pieces of evidence, due to dependencies. We propose to report evidential support for separate pieces of evidence, as well as one combined evidential support for the entire set of evidence. Finally, we also propose to point out specifically which evidence distinguishes between scenarios and which does not (it gives equal support to more than one scenario).

To report the amount of evidential support or attack by a piece (or a set) of evidence, we propose to use a verbal scale to translate the evidential support as calculated with the fraction above. Similar scales are often used for the reporting of likelihood ratios, since likelihood ratios (and arguably our measure of evidential support as well) can be hard to interpret for a judge or jury (Nordgaard et al. 2012). Our proposed verbal scale is shown in Table 1. This scale was adapted from the standard quantitative scale published by the Association of Forensic Science Providers (2009). The original scale was intended for likelihood ratios associated

Table 1 A qualitative scale for evidential support

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

with forensic evidence such as DNA matching, so to accommodate for our use of the scale with evidential support, we adapted the scale to work with slightly less extreme numbers (which are usually associated with DNA evidence) and to include translations below 1 (which were not included on the original scale).

4.3 Scenario quality

In Sect. 3.2, definitions were given for the three factors that determine a scenario's quality: completeness, consistency and plausibility. A good scenario should be complete, consistent and plausible. Completeness and consistency were defined as boolean notions, while plausibility is can be gradual. As soon as a scenario is incomplete it should not be considered as an acceptable alternative. Therefore, we proposed that only complete scenarios should be modelled in the network. Similarly, when a scenario is inconsistent it should not be an acceptable alternative. This is achieved by the way we modelled inconsistency in the network, such that the probability of the scenario node becomes 0 as soon as there is an inconsistency within the scenario. In contrast with completeness and consistency, when a scenario is implausible it is not immediately dismissed. With sufficient support, even a very implausible scenario can become credible (Wagenaar et al. 1993). This makes plausibility a relevant factor to report in an explanation of the network.

In fact, we observe that the effect of evidential support for an implausible element can be much stronger than the effect of evidential support for a plausible element. As an example, consider the following scenario: 'Jane and Mark had a fight. Jane had a knife, and she threatened Mark with it. Mark hit Jane, who then dropped the knife. Mark fell on the knife and died.' This is an implausible scenario, particularly because of the implausible element that Mark fell on the knife, and the implausible connection that Mark died from falling on the knife. But with sufficient support for these implausible parts, the scenario can go from very improbable to quite probable. On the other hand, evidence that Jane had a knife to begin with has much less effect, since the event itself is already credible enough.

From the three factors above, plausibility is thus of interest for an explanation of the network, and in particular for understanding how the posterior probability of a scenario changed as result of the evidence. We furthermore observe that implausible elements are worth reporting when they are supported by evidence to become (much) more probable, but also when they remain unsupported. The latter point out so-called *evidential gaps*: implausible elements of the scenario that remain unsupported.

In what follows, we propose to report implausible elements or connections using a threshold of 0.01. That is to say, whenever $\Pr(V = \mathtt{T}) \leq 0.01$, we consider element V implausible, and whenever $\Pr(V_2 = \mathtt{T}|V_1 = \mathtt{T}) \leq 0.01$ we consider the connection between V_1 and V_2 as implausible. In general, we say that V_j is implausible given parents $V_1, \ldots, V_i$ when $\Pr(V_j = \mathtt{T}|V_1 = \mathtt{T}, \ldots, V_i = \mathtt{T}) \leq 0.01$.

4.4 A reporting format

With the material from the previous subsections, information about the scenarios in a Bayesian network can be reported. For each scenario, we propose to report the following:

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

Consider our fictional case about Jane and Mark once again, with a scenario about Jane stabbing Mark and a scenario about Mark falling on a knife. Suppose we also have the following evidence: Jane had a knife (Jane's testimony), Mark's dead body, and stab wounds which better fit the scenario about Jane stabbing Mark. The graphical structure for this fictional example was made using our construction method from Vlek et al. (2014) enhanced with scenario schemes, as is shown in Fig. 7. Probabilities were also assessed, though these are clearly very subjective and only made explicit for the purpose of illustrating the proposed reporting format. A full model for this case is available at www.charlottevlek.nl/networks. A report of this network could then resemble the following:

– *Scenarios in the network*:

 – Scenario 1 (prior probability: 0.001, posterior probability: 0.97): Jane and Mark had a fight and Jane had a knife. Then Jane stabbed Mark. Therefore, Mark died.
 – Scenario 2 (prior probability: 0.001, posterior probability: 0.01): Jane and Mark had a fight and Jane had a knife. Then Jane dropped the knife. Then Mark fell on the knife. Therefore, Mark died.

– *Evidence related to each scenario*

 – Evidence for and against scenario 1:

 • Jane's testimony: Jane had a knife: moderate evidence to support scenario 1.
 • Dead body of Mark: strong evidence to support scenario 1.
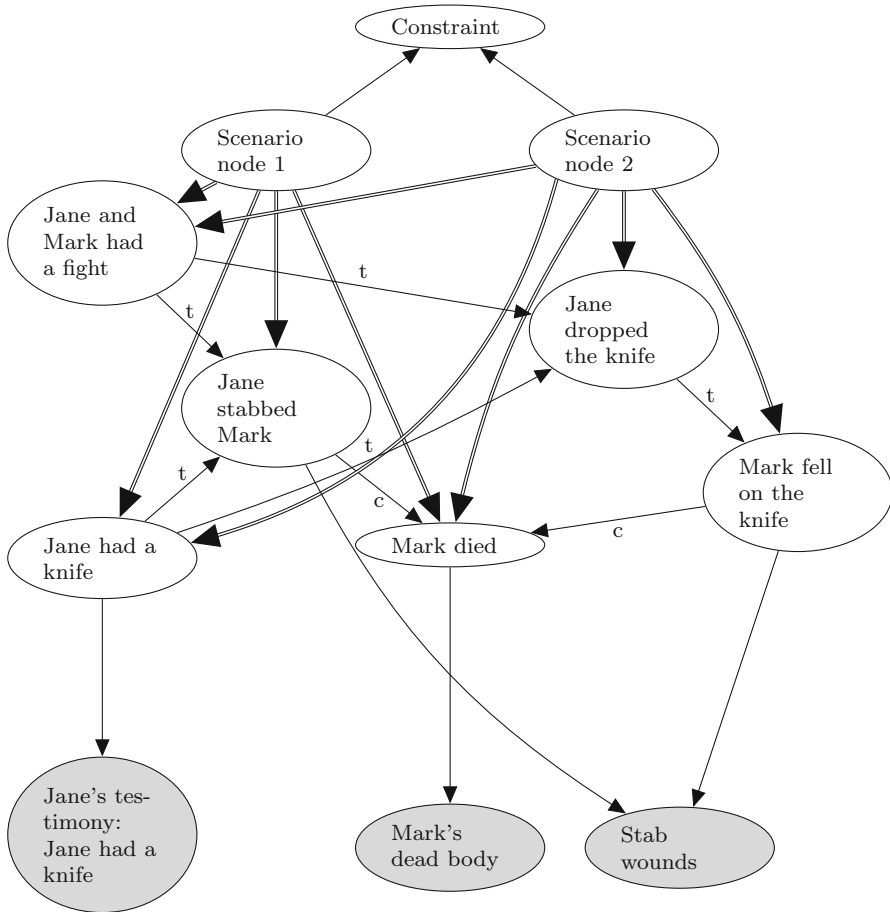
Fig. 7 A network for the example case about Jane and Mark. Evidential nodes are indicated as grey nodes

- Stab wound of Mark: strong evidence to support scenario 1.
- *Combined strength of evidence*: strong evidence to support scenario 1.

– Evidence for and against scenario 2:

- Jane's testimony: Jane had a pocket knife: moderate evidence to support scenario 2.
- Dead body of Mark: strong evidence to support scenario 2.
- Stab wound of Mark: weak evidence to support scenario 2.
- *Combined strength of evidence*: weak evidence to support scenario 2.

- – *Distinguishing evidence*: Stab wound of Mark is the only distinguishing evidence, the other evidence is neutral.

- – *Implausible elements/connections in each scenario*

    - – Scenario 1 contains the implausible element `Jane had a knife`, which is supported with evidence.
    - – Scenario 2 contains the implausible element `Mark fell on the knife` given `Jane dropped the knife`. Given the evidence, the event `Mark fell on the knife` remains implausible, it is thus an evidential gap.
    - – Scenario 2 contains the implausible element `Mark died` given `Mark fell on the knife`. Given the evidence this remains implausible, it is thus an evidential gap.

Such a report can potentially be produced automatically. A scenario in text form can be produced with the method from Sect. 4.1.2, and the prior and posterior probability of the scenario can be read from the network (at the scenario node corresponding to that scenario).

For each piece of evidence, the evidential support of that evidence for the scenario can be calculated with the measure from Sect. 4.2, and translated to a verbal account with the verbal scale from that same section. Similarly, a combined measure of evidential support is calculated and reported verbally. Finally, when a report is made for multiple scenarios, neutral evidence (with equal evidential support for both scenarios) is pointed out.

Implausible elements and connections can be found in the network as elements or connections with a probability below 0.01, conditioned on their parents within the scenario as explained in Sect. 4.3. When an implausible element has no evidence directly connected to it, it is reported as an evidential gap. When there is evidence for it, it is pointed out as an implausible element with support.

5 Case study

In this section, we perform a case study to test the proposed explanation method. The Bayesian network is constructed using our previously proposed construction method from Vlek et al. (2014), with the adaptation proposed in Sect. 3. In the case, a suspect is thought to have helped to move a dead body to a location out of town after someone else committed the killing in a house in town. The suspect was convicted of this, but later the Court of Appeals concluded that there was not enough evidence to convict, so the suspect was found innocent. The appeal case can be found (in Dutch) on www.rechtspraak.nl using code ECLI:NL:GHARL:2014:8941.

This case is of interest for testing our method because it includes some probabilistic reasoning as well as several alternative scenarios about what may have happened. The probabilistic reasoning (though non-numerical) is concerned with DNA-evidence in this case. DNA traces matching with the suspect were found on the victim's body. Several alternative activity-level hypotheses were compared: the

suspect moved the body, thereby leaving traces, or the suspect was not involved in moving the body but his DNA was transferred from a couch in the victim's home. In the appeal case, another hypothesis was added, namely that the suspect was not involved in moving the body, but his DNA was transferred from a blanket (used by the suspect some time prior to the killing), which was used to wrap the body and transport it to another location. For each of these hypotheses, the likelihood of finding DNA traces given each hypothesis was compared qualitatively in the appeal case. In our analysis of the case, we will provide a numerical interpretation of this qualitative discussion in our model.

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

5.1 Scenario 1

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

A scenario scheme for this scenario is shown in Fig. 8 (left), as well as the scenario as modelled using this scheme (right). In this specific case, details about how they moved the body are of interest. The element 'Adam and Bert moved Chris's body' is thus unfolded to a subscenario, as shown in Fig. 9. Furthermore, the element 'Adam, Bert and Chris were involved in a cannabis operation' is also unfolded, since it will become important in relation to the other scenarios (see below).
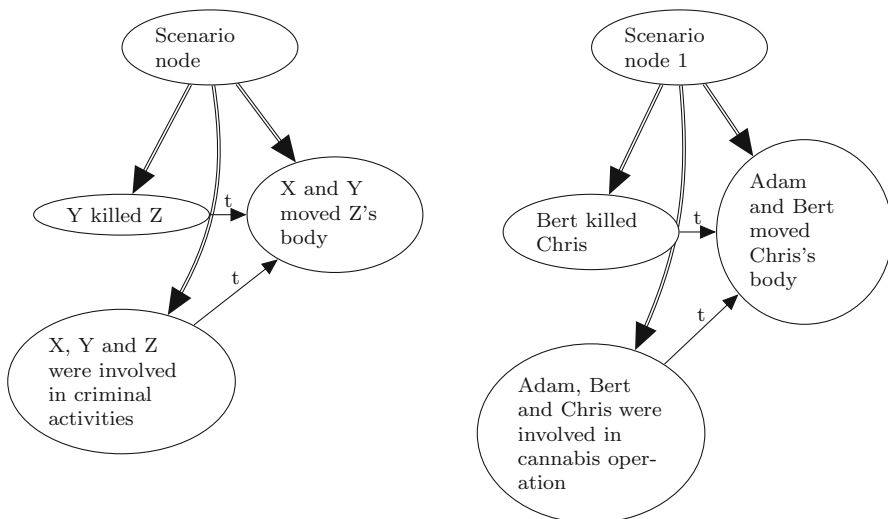
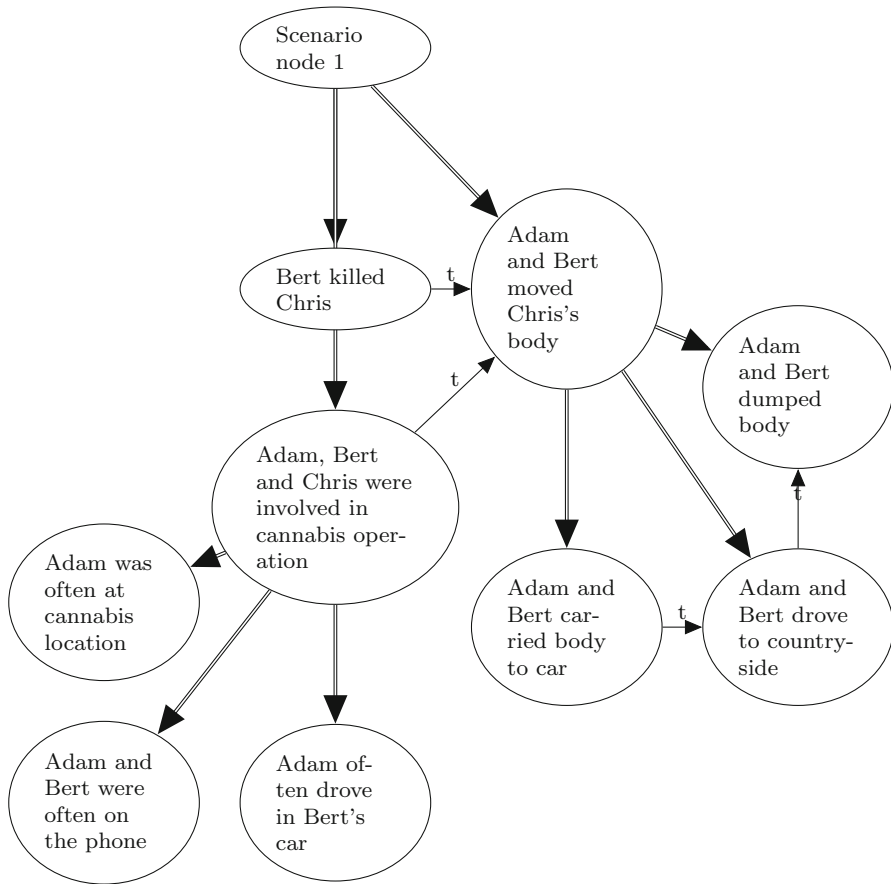**Fig. 8** A scenario scheme idiom (*left*) and the initial scenario structure for scenario 1 (*right*)

Fig. 9 The first scenario with subscenarios

5.2 Scenario 2

In the second scenario, Bert worked alone and also moved the body by himself. Note that this is effectively an innocent scenario for the current suspect Adam. This scenario can be represented with a scenario scheme with the same structure as the one that was used earlier. This second scenario also includes the element 'Adam, Bert and Chris were involved in a cannabis operation', because this element (when worked out as a subscenario as shown in Fig. 9) helps to explain why evidence was found that related the crime to Adam: DNA traces of Adam were on the couch at the cannabis plantation and then transferred to Chris's body. A network with both scenarios is shown in Fig. 10.
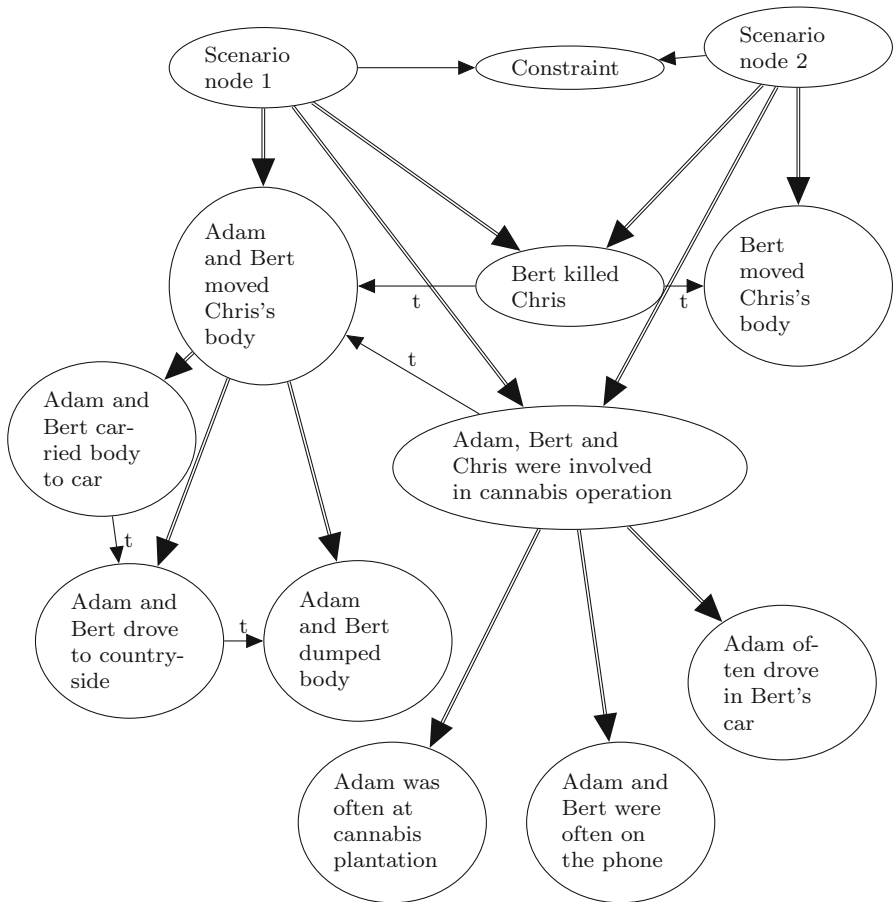
Fig. 10 The first two scenarios in one network

5.3 Scenario 3

In the appeal case, a third scenario was considered in addition to the two scenarios described above. In this scenario, Bert also worked alone and moved the body by himself, but now he used a blanket to move the body, and this blanket contained traces of Adam, because Adam sometimes stayed at the cannabis plantation during the night. As we will see later on, the probability of a transfer of Adam's DNA is now higher than when transfer took place via the couch. This third scenario is represented with the same scenario scheme as the previous (second) scenario. A network with all three scenarios is shown in Fig. 11, constructed with the method from Vlek et al. (2014), adapted with the use of scenario schemes and scenario quality.
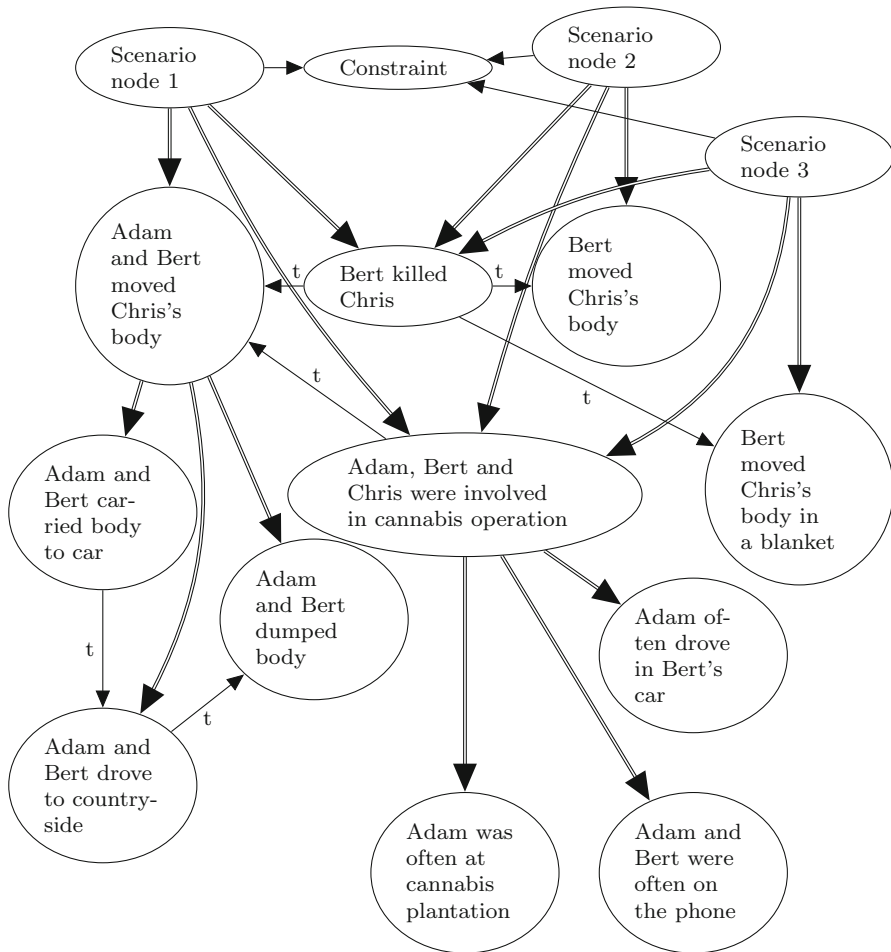
Fig. 11 The three scenarios in one network

5.4 The evidence

Evidence in this case is the following:

- Adam's car was not seen on any surveillance camera's (Adam's car on ARS cameras)
- A DNA match was found between Adam's profile and traces on Chris's body (DNA match)
- A hair matching Adam's was found on some duct tape (Hair on duct tape)
- Bert was already convicted for the killing of Chris (Bert's conviction)
- Chris's body was found in the countryside (Body in countryside)
- Adam and Bert talked on the phone several times at the night of the killing (Phone calls Adam and Bert)
- Traces in Bert's car left by Adam (Traces of Adam in car)

For each piece of evidence, an evidential node is added to the network. This leads to a graph as shown in Fig. 12 in the appendix.

5.5 The numbers

A probabilistic discussion of the DNA match played an important role in the appeal case. We capture the qualitative analysis in the numbers of the probability table for the node DNA match. In the initial case, the court considered only two activity-level hypotheses: Adam and Bert carried the body together, or Bert moved Chris's body by himself and Adam's DNA was transferred some other way (because Adam was often at the cannabis plantation). An expert stated that the DNA evidence would be 'much more likely' to be found under the hypothesis that Adam and Bert carried the body than under the hypothesis that Bert moved Chris's body by himself and the DNA was transferred some other way. In the appeal case, a third hypothesis was considered, namely that Bert moved the body in a blanket which Adam had used previously when he was sleeping at the cannabis plantation. From the court records it becomes apparent that given this third hypothesis, the DNA evidence is still considered quite likely, although it is unclear how it compares to the other two hypotheses. In the probability table of DNA match, these considerations can be quantified with some interpretation. We set the probabilities as shown in Table 2, such that

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

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

For the assessment of other numbers, elicitation techniques exist, see, e.g., Renooij (2001). The scenario nodes are a special case, since there the prior probability is expressed that each scenario as a whole is true. For lack of further information about this, we propose to set the probability of each scenario to be true to the low probability of 0.001, well below our threshold of 0.01 for plausibility.

Table 2 The probability table for `DNA match`

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

Furthermore, numbers in the probability tables of the network in Fig. 12 are partially constrained due to the scenario scheme idioms. The probability of an element of a scenario conditioned on the scenario node being true is always one: $\Pr(\text{Element} = \texttt{True}|\text{Scenarionode} = \text{T}) = 1$, and similarly for elements of a subscenario (such as $\Pr(\text{Adam and Bert carried body to car} = \text{T}|\text{Adam and Bert moved Chris's body} = \text{T}) = 1$).

The assessment of probabilities for elements of the scenarios needs to take into account plausibility, as explained in Sect. 3.2. For example, the element that Bert killed Chris seems implausible, and we would like to set the numbers such that $\Pr(\text{Bert killed Chris} = \text{T}) = 0.01$. To calculate the numbers as they should appear in the probability table, we use

$$
\begin{aligned}
0.01 =& \Pr(\text{Bert killed Chris} = \text{T})\\
=& \Pr(\text{Bert killed Chris} = \text{T}|\text{ScN1} = \text{T}) \cdot \Pr(\text{ScN1} = \text{T})\\
& + \Pr(\text{Bert killed Chris} = \text{T}|\text{ScN1} = \text{F}) \cdot \Pr(\text{ScN1} = \text{F})\\
=& 1 \cdot \Pr(\text{ScN1} = \text{T})\\
& + \Pr(\text{Bert killed Chris} = \text{T}|\text{ScN1} = \text{F}) \cdot \Pr(\text{ScN1} = \text{F})\\
=& 1 \cdot 0.001 + \Pr(\text{Bert killed Chris} = \text{T}|\text{ScN1} = \text{F}) \cdot 0.999
\end{aligned}
$$

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

$$
\Pr(\text{Bert killed Chris} = \text{T}|\text{ScN1} = \text{F}) = \frac{0.01 - 0.001}{0.999} \approx 0.009.
$$

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

$$\Pr(\text{Adam and Bert moved Chris's body} = \text{T}|\text{Bert killed Chris} = \text{T},$$
$$\text{Adam, Bert and Chris were involved in cannabis operation} = \text{T})$$
$$= 0.1$$

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

5.6 The report

With a full Bayesian network available, a report can be compiled to explain the network. This report will consist of three parts, as discussed in Sect. 4.4:

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

It is possible to extract this report fully automatically, since information needed for all three parts can be computed from the network directly. For illustration purposes, a report was compiled manually for the network in Fig. 12.

- *Scenarios in the network*:
    - Scenario 1 (prior probability: 0.001, posterior probability: 0.5296): Bert killed Chris, and [Adam, Bert and Chris were involved in cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car]. Then [Adam and Bert moved Chris's body: Adam and Bert carried body to car. Then Adam and Bert drove to countryside. Then Adam and Bert dumped body].
    - Scenario 2 (prior probability: 0.001, posterior probability: 0.1180): Bert killed Chris, and [Adam, Bert and Chris were involved in cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car]. Then Bert moved Chris's body.
    - Scenario 3 (prior probability: 0.001, posterior probability: 0.2913): Bert killed Chris, and [Adam, Bert and Chris were involved in cannabis operation: Adam was often at cannabis location and Adam and Bert were often on the phone and Adam often drove in Bert's car]. Then Bert moved Chris's body in a blanket.

– *Evidence related to each scenario*

  – Evidence for and against scenario 1:

  - `Adam's car on ARS cameras = F`: weak evidence to attack scenario 1
  - `DNA match = T`: moderate evidence to support scenario 1
  - `Hair on duct tape = T`: moderate evidence to support scenario 1
  - `Bert's conviction = T`: moderate evidence to support scenario 1
  - `Body in countryside = T`: strong evidence to support scenario 1
  - `Phone calls Adam and Bert = T`: weak evidence to support scenario 1
  - `Traces of Adam in car = T`: weak evidence to support scenario 1
  - *Combined strength of evidence*: strong evidence to support scenario 1

  – Evidence for and against scenario 2:

  - `Adam's car on ARS cameras = F`: weak evidence to attack scenario 2
  - `DNA match = T`: moderate evidence to support scenario 2
  - `Hair on duct tape = T`: moderate evidence to support scenario 2
  - `Bert's conviction = T`: moderate evidence to support scenario 2
  - `Body in countryside = T`: strong evidence to support scenario 2
  - `Phone calls Adam and Bert = T`: weak evidence to support scenario 2
  - `Traces of Adam in car = T`: weak evidence to support scenario 2
  - *Combined strength of evidence*: strong evidence to support scenario 2

  – Evidence for and against scenario 3:

  - `Adam's car on ARS cameras = F`: weak evidence to attack scenario 3
  - `DNA match = T`: moderate evidence to support scenario 3
  - `Hair on duct tape = T`: moderate evidence to support scenario 3
  - `Bert's conviction = T`: moderate evidence to support scenario 3
  - `Body in countryside = T`: strong evidence to support scenario 3
  - `Phone calls Adam and Bert = T`: weak evidence to support scenario 3
  - `Traces of Adam in car = T`: weak evidence to support scenario 3
  - *Combined strength of evidence*: strong evidence to support scenario 3

  – *Distinguishing evidence*: All evidence is distinguishing evidence, except `Bert's conviction = T` and `Body in countryside = T`.

– *Implausible elements/connections in each scenario*

  – All scenarios contain the implausible element `Bert killed Chris`. This element is supported by `Bert's conviction = T`.

5.7 Evaluation

With the proposed explanation method, our goal was to provide a juror with insight about the model in terms of narrative properties such as scenario quality and evidential support. To this end, we also proposed an adaptation of the construction method in Vlek et al. (2014). In this section, we use the case study to evaluate the adapted construction method and the explanation method using the three questions from the introduction:

1. Does the use of scenario schemes in the adapted construction method assist the modeller when constructing a Bayesian network?
2. Are various scenarios and their properties (e.g., quality) adequately represented in the Bayesian network?
3. Does the report contain the required information about scenarios, scenario quality and evidential support?

Regarding the first criterion, the use of scenario schemes provides structure to the construction process, which should simplify the task of a modeller. Ultimately, the idea is to have a database of scenario schemes at the disposal of a modeller. When modelling the case, no such database was available (yet). Nonetheless, we found that the sheer task of working out the abstract structure of a scenario scheme before modelling the scenario itself is helpful to systematise the construction process. For example, by working out a scenario scheme for the initial first scenario (Bert killed Chris; Adam, Bert and Chris were involved in a cannabis operation; Adam and Bert moved Chris's body), we found that the details about moving the body did not belong in the main scenario, but rather in a subscenario.

Regarding the second criterion, the quality of a scenario is represented in a network in terms of completeness, consistency and plausibility. In this case, all scenarios were complete. If we were to encounter an incomplete scenario (e.g., in which Chris's body was moved but it is not described who killed him), a scenario scheme would demand additional elements of a scenario to be specified, such that it can be modelled as a complete scenario. This case included only consistent scenarios. If an inconsistency were encountered, it could be represented with a constraint on the inconsistent elements. Finally, all three scenarios in this case contained the implausible element 'Bert killed Chris'. This implausibility was represented in the network with appropriate probabilities for the node `Bert killed Chris` such that the probability of this event occurring is low. It is thus possible to represent scenario quality adequately in the network. Representing a scenario is nonetheless still a difficult task due to all the probabilities that need to be elicited. While elicitation techniques exist, it is not yet clear how well these work with our current method.

As for the third criterion, the report provides information about which scenarios are in the network, their prior and posterior probability, their quality and the evidential support for each scenario. In the case study, a report was compiled containing all the required information. A possibly surprising result is that the DNA evidence is reported to be (only) 'moderate evidence to support'. This is because rather than reporting the support

of the DNA match for the hypothesis that the DNA belonged to Adam (which we expect to be very high), what is reported is the evidential support related to activity-level hypotheses about how the DNA traces came to be. We furthermore observe that the report supports the decision as made by the Court of Appeals. After an initial conviction, the appeal case was decided as unconvincing, and the suspect was found innocent. This conclusion, that the case was unconvincing, is reflected in the report. First of all, in the report we see that the first scenario has a higher posterior probability than the second and third scenario, but not so high that it is beyond reasonable doubt. In fact, the combination of evidence is reported as 'strong evidence to support' for all three scenarios. This is because the evidential support for each of the three scenarios was in the same order of magnitude (between 100 and 1000, see Table 1) Secondly, the third scenario is more probable than the second scenario, which supports the choice of the court to also include this third scenario. Thirdly, the report reveals that all scenarios are of similar quality, all three containing the same implausible element which is supported by evidence (Bert's conviction). Finally, all pieces of evidence that support scenario 1 also support scenario 2 and 3, though to a lesser extent. And while most of the evidence distinguishes between the three scenarios, none of the scenarios can be ruled out.

6 Related work

Kadane and Schum ([1996](#)) were the first to analyse an entire legal case in terms of probabilities. Since then, Bayesian networks have become popular for modelling parts of a case (see, e.g., Taroni et al. [2006](#)), but much less so for representing entire cases. Nonetheless, some work has been done to simplify the construction of a complex network for a legal case (i.e., a network that is larger than a few nodes for analysing a technical result) by using basic structures that recur throughout various cases as building blocks. This was proposed by Hepler et al. ([2004](#)), who used object-oriented Bayesian networks. Fenton et al. ([2013](#)) proposed a list of legal idioms which served as basic building blocks for legal cases. In previous work, we extended these idioms with narrative idioms for representing scenarios in a Bayesian network (Vlek et al. [2014](#)), and in this paper we proposed the use of scenario scheme idioms. These idioms not only make it possible to represent scenarios in a Bayesian network, they also provide more context-specific structure than our previous narrative idioms, which means less work for the modeller. Finally, the scenario scheme idioms allow for an interpretation of the narrative concept of completeness in a Bayesian network context.

Using the scenario scheme idioms, the network is constructed in such a way that scenarios are clearly represented, which makes it easier to understand the network. In particular, as ingredients for a report, we proposed a method to extract scenarios from the network, and we proposed to report about evidential support and scenario quality, in particular the plausibility of the scenario. The fact that our explanation method requires that the Bayesian network is constructed with a specific method can be viewed as an advantage as well as a disadvantage: while the explanation method will not work on just any Bayesian network, integrating the construction and explanation methods provides one single approach for constructing and explaining a Bayesian network with scenarios.

An overview of research on explaining Bayesian networks can be found in Lacave and Díez (2002). Typically, the explanation of Bayesian networks attempts to explain why certain modelling choices were made and why the network produces a certain result given these choices. The particular combination with a narrative approach for explaining a network has not been made before, although there is some work by Druzdzel (1990), who uses scenarios to explain Bayesian networks. However, the scenarios used in Druzdzel (1990) consists of configurations of nodes and satisfy no further narrative properties such as completeness, consistency and plausibility. In particular, in Druzdzel (1990), the following would be considered a scenario, while in our work it is not: `Jane had a knife = F`, `Jane stabbed Mark = F`, `Mark died = T`. Other work specifically related to explaining Bayesian networks for legal cases is by Timmer et al. (2014, (2015) and Keppens (2012), who both extract arguments from Bayesian networks. In particular, Timmer et al. (2014) also uses a similar measure to ours to calculate evidential support.

In our combined approach using Bayesian networks and narratives, our goal was to have a formal model of the case in a Bayesian network, while allowing a judge or juror to reason in a way that is more intuitive for them, with the use of scenarios. By reporting about the scenarios in the network, evidential support and scenario quality, we aimed to report sufficient information such that a juror could apply a narrative approach to reasoning with legal evidence. To this end, we provided interpretations of the narrative properties of completeness, consistency and plausibility as they are used by Pennington and Hastie (1993) and Bex (2011). With the use of scenario scheme idioms, our interpretation of completeness remains close to its interpretation by Pennington and Hastie (1993) and Bex (2011). Consistency is also interpreted similarly to Pennington and Hastie (1993) and Bex (2011): a scenario is either consistent or inconsistent (which leads to a probability of 0 for the scenario node). However, we remark here that a probabilistic approach would also lend itself for a gradual interpretation of consistency, in which a scenario could be consistent to a certain extent. For example, a conflict measure as discussed in Jensen and Nielsen (2007) could be adapted to serve as a measure of consistency within a scenario. In this paper we chose to represent consistency such that it resembles the concept of consistency by Pennington and Hastie (1993) and Bex (2011). Finally, Pennington and Hastie (1993) and Bex (2011) relate plausibility to common sense knowledge about the world. We argued that in a Bayesian network, common sense knowledge is represented in the probability tables, which express how probable an event is without having any further information than our basic common sense. Plausibility can thus be read from a Bayesian network when no evidence has been entered in the network yet.

Our attempt to combine a narrative approach with a probabilistic approach is led by a practical goal, namely to make Bayesian networks more insightful for a judge or jury, but also by the more fundamental goal to understand the relations between these approaches. Typically, three distinct approaches are used to formalize reasoning with legal evidence: the narrative approach, the probabilistic approach and the argumentative approach. Combinations of these approaches have been studied by Bex (2011), who combined arguments and scenarios, by Timmer et al. (2015), who considered Bayesian networks and arguments, and by Verheij (2014), who combined arguments, scenarios and probabilities. An overview of these

combined approaches can be found in Verheij et al. (2015). Our contribution with this paper is to combine Bayesian networks and scenarios such that scenario-based reasoning can support the explanation of Bayesian networks.

7 Conclusion

In this paper, an explanation method for understanding Bayesian networks with scenarios was proposed. To this end, a previously proposed method for constructing a Bayesian network with scenarios was adapted for the purpose of explaining a network built with this construction method. A probabilistic interpretation of scenario quality was provided in terms of completeness, consistency and plausibility. The original construction method was adapted with the use of scenario schemes, which structure and simplify the construction process, and scenario quality was represented in the network. From the resulting structure, scenarios in text form could be extracted and presented in a report which also included evidential support and information about the quality of the scenario.

The proposed method combines a probabilistic approach with a narrative approach to reasoning with legal evidence. The advantage of using a Bayesian network is that statistical evidence can be analysed in relation to the entire case. The advantage of a narrative approach is that it provides the global perspective on the case as a whole, and that it provides an insightful way for a judge or jury to reason about the case.

A case study was performed to evaluate the proposed method. We concluded that (1) the adapted construction method is helpful for a modeller when constructing a network, (2) scenario quality can now be represented in a network, although representing a scenario in a network is still not an easy task due to all the probabilities that need to be elicited and (3) scenarios, scenario quality and evidential support are conveyed in the proposed report.

In future research, it will be of interest to find whether the proposed report (perhaps in a more polished appearance) is insightful for a judge or jury. Other future research could include a combination of methods in which a full probabilistic analysis is not required. Currently, our proposed method requires a fully specified Bayesian network for the case, to allow a global perspective. In future research it would be desirable to combine Bayesian networks and narrative in such a way that statistical information can be analysed with a probabilistic approach, while non-statistical information does not require a probabilistic analysis.
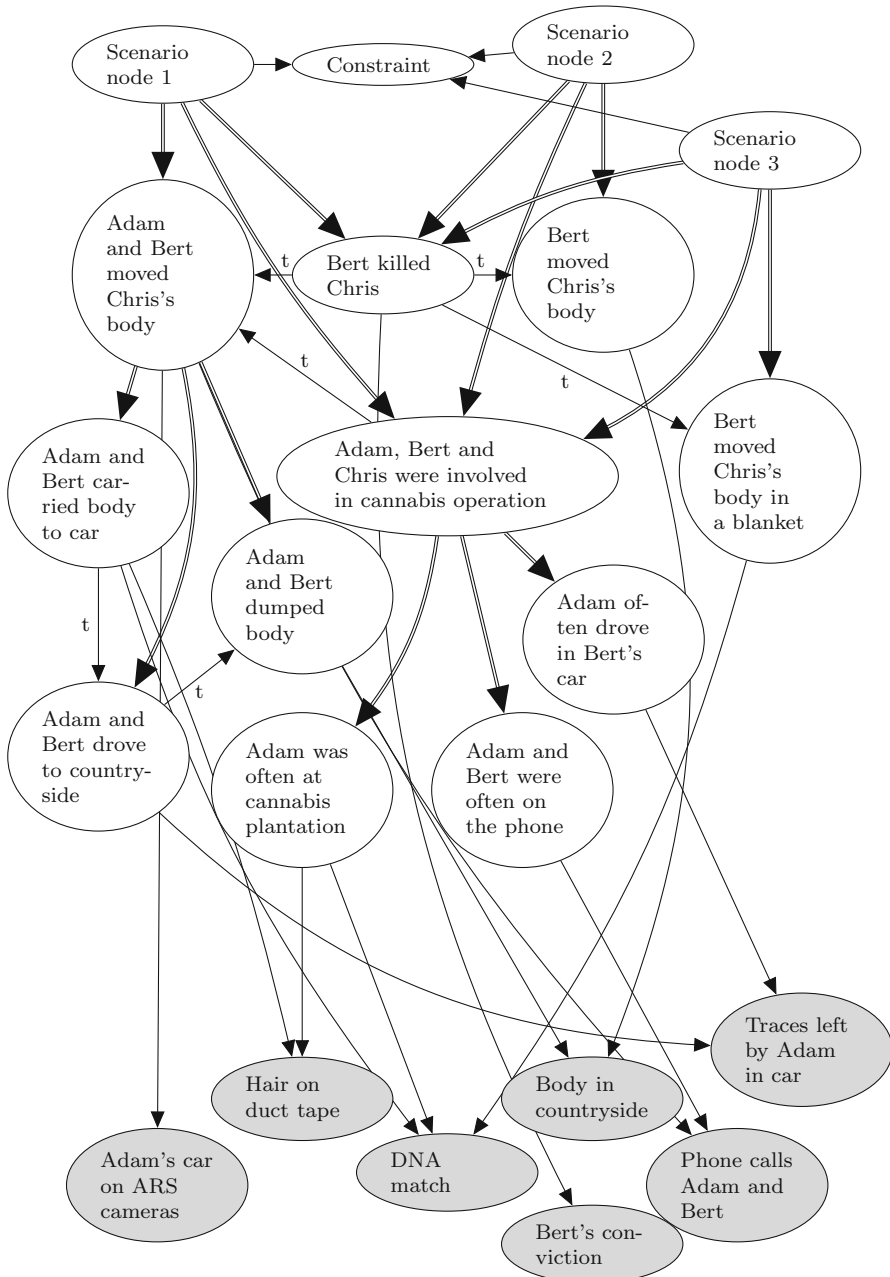
Appendix

See Fig. 12.

Fig. 12 A network for the case study: The three scenarios with evidence. Evidential nodes are indicated as grey nodes

References

Association of Forensic Science Providers (2009) Standards for the formulation of evaluative forensic science expert opinion. Sci Justice 49(3):161–164

Bennett W, Feldman M (1981) Reconstructing reality in the courtroom: justice and judgment in American culture. Methuen-Tavistock, London

Bex F (2011) Arguments, stories and criminal evidence, a formal hybrid theory. Springer, Dordrecht

Dawid A (2009) Beware of the DAG. J Mach Learn Res Workshop Conf Proc 6:59–86

Druzdzel M (1990) Using scenarios to explain probabilistic inference. In: Eorking notes of the AAAI-90 Workshop on Explanation, pp 133–141

Fenton N, Neil M (2011) Avoiding probabilistic reasoning fallacies in legal practice using Bayesian networks. Aust J Leg Philos 36:114–151

Fenton N, Neil M, Lagnado D (2011) Modelling mutually exclusive causes in Bayesian networks. Available online: www.eecs.qmul.ac.uk/~norman/papers/mutual_IEEE_format_version

Fenton N, Neil M, Lagnado D (2013) A general structure for legal arguments using Bayesian networks. Cogn Sci 37:61–102

Fenton N, Berger D, Lagnado D, Neil M, Hsu A (2014) When 'neutral' evidence still has probative value (with implications from the Barry George Case). Sci Justice 54(4):274–287

Hepler A, Dawid A, Leucari V (2004) Object-oriented graphical representations of complex patterns of evidence. Law Probab Risk 6:275–293

Jensen F, Nielsen T (2007) Bayesian networks and decision graphs. Springer, New York

Kadane J, Schum D (1996) A probabilistic analysis of the Sacco and Vanzetti evidence. Wiley, New York

Keppens J (2012) Argument diagram extraction from evidential Bayesian networks. Artif Intell Law 20(2):109–143

Lacave C, Díez F (2002) A review of explanation methods for Bayesian networks. Knowl Eng Rev 17(2):107–127

Nordgaard A, Ansell R, Drotz W, Jaeger L (2012) Scale of conclusions for the value of evidence. Law Probab Risk 11(1):1–24

Pennington N, Hastie R (1993) Inside the juror. Cambridge University Press, Cambridge, chap The Story Model for Juror Decision Making, pp 192–221

Renooij S (2001) Probability elicitation for belief networks: issues to consider. Knowl Eng Rev 16(3):255–269

Schank R, Abelson R (1977) Scripts, plans, goals and understanding, an inquiry into human knowledge structures. Lawrence Erlbaum, Hillsdale

Taroni F, Aitken C, Garbolino P, Biedermann A (2006) Bayesian networks and probabilistic inference in forensic science. Wiley, Chichester

Timmer S, Meyer J, Prakken H, Renooij S, Verheij B (2014) Extracting legal arguments from forensic Bayesian networks. In: Hoekstra R (ed) The 27th international conference on legal knowledge and information systems (JURIX) 2014. IOS Press, Amsterdam, pp 71–80

Timmer S, Meyer J, Prakken H, Renooij S, Verheij B (2015) A structure-guided approach to capturing Bayesian reasoning about legal evidence in argumentation. In: Proceedings of the 15th international conference on artificial intelligence and law. ACM Press, New York, pp 109–118

Verheij B (2014) To catch a thief with and without numbers: arguments, scenarios and probabilities. Law Probab Risk 13(3–4):307–325

Verheij B, Bex FJ, Timmer ST, Vlek CS, Meyer J-J, Renooij S, Prakken H (2015) Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning. Law Probab Risk 15(1):35–70

Vlek C, Prakken H, Renooij S, Verheij B (2014) Building Bayesian networks for legal evidence with narratives: a case study evaluation. Artif Intell Law 22(4):375–421

Vlek C, Prakken H, Renooij S, Verheij B (2015a) Constructing and understanding Bayesian networks for legal evidence with scenario schemes. In: Proceedings of the 15th international conference on artificial intelligence and law. ACM Press, New York, pp 128–137

Vlek C, Prakken H, Renooij S, Verheij B (2015b) Representing the quality of crime scenarios in a Bayesian network. In: Rotolo A (ed) The 28th international conference on legal knowledge and information systems (JURIX) 2015, IOS Press, Amsterdam (in press)

Wagenaar W, van Koppen P, Crombag H (1993) Anchored narratives: the psychology of criminal evidence. Harvester Wheatsheaf, Hemel Hempstead

Wellman P, Henrion M (1993) Explaining 'explaining away'. IEEE Trans Pattern Anal Mach Intell 15(3):287–292