Scenario-Based Bayesian Networks for Legal Evidence

Dahlman, Christian; Jellema, Hylke

Published in:
Artificial Intelligence and Law

2026

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

Artificial Intelligence and Law
https://doi.org/10.1007/s10506-026-09522-9

ORIGINAL RESEARCH

Scenario-based Bayesian networks for legal evidence

Christian Dahlman[1] · **Hylke Jellema**[2,3]

Received: 11 September 2025 / Accepted: 27 April 2026

Abstract
The paper explores how the construction of Bayesian networks for legal evidence can be assisted by scenario models, especially how *spatio-temporal constraints* can be transferred from scenario models to Bayesian models. The paper introduces two new methods for constructing scenario-based Bayesian networks: *non-Boolean time-nodes with event-states* and *non-Boolean event-nodes with time-states*.

Keywords Law · Evidence · Bayesian network · Scenario

1 Introduction

Bayesian networks and scenario analysis are two important approaches to legal evidence that have been developed separately for several decades. In this paper, we will show how they can be combined. We will demonstrate how the construction of Bayesian networks from the perspective of the legal fact-finder (judge or jury) can be helped by scenario modeling. In particular, we will explore how scenario models can be of assistance in the modeling of *spatio-temporal constraints* in Bayesian networks.

In a previous attempt to combine the two approaches, it has been suggested that so-called *scenario-nodes* can be employed to incorporate scenarios in Bayesian net-

Christian Dahlman and Hylke Jellema have contributed equally to this paper (the order of names is purely alphabetical). Thanks to Leya Hampson, David Lagnado, Ludi van Leeuwen, Moa Lidén, Anne Ruth Mackor and Elaine Strittmatter for helpful comments on previous versions of this manuscript.

✉ Christian Dahlman
christian.dahlman@jur.lu.se

Hylke Jellema
h.jellema@rug.nl

[1] Faculty of Law, Lund University, Lund, Sweden

[2] University of Groningen, Groningen, Netherlands

[3] Utrecht University, Utrecht, Netherlands

Published online: 23 May 2026

works (Vlek et al 2014). However, as we will show, this modeling method leads to mathematically incorrect propagations (Sect. 3).

In this paper, we will propose two methods for modeling scenarios in Bayesian networks that have not previously been explored in the literature on legal evidence: *time-nodes with event-states* (Sect. 4) and *event-nodes with time-states* (Sect. 5). As we shall see, they bring scenario modeling into Bayesian networks in a clear and simple way that helps the fact-finder take account of the spatio-temporal constraints that arise when scenario-elements are incompatible.

2 Bayesian networks and scenario analysis

Bayesian networks are graphs that model the probabilistic dependencies between the hypotheses at issue in a case and the observed facts that are considered as evidence. Hypotheses that are connected to each other create structures where a certain piece of evidence directly affects some hypothesis and indirectly affects other hypotheses in the network. The model therefore helps to assess how the probabilities at issue are affected by the evidence, given certain assumptions about the dependencies. For a formal mathematical definition of Bayesian networks, see the Appendix.

Bayesian networks are used for probability assessments in many different contexts (cf. Fenton & Neil 2019). In the legal context, two kinds of Bayesian networks have been developed: *forensic expert models* and *fact-finder models*. Forensic expert models help forensic experts answer the questions that are put to them in their role of expert witnesses. A forensic expert model is restricted to a forensic aspect of the case and helps the expert assess how strongly the forensic evidence supports a particular hypothesis (cf. Taroni et al 2014), for example how strongly a DNA match supports the hypothesis that the defendant is the source of a DNA trace found at the crime scene. Fact-finder models are constructed from the perspective of the legal fact-finder (judge or jury) and offer assistance in assessing the probability of the ultimate hypothesis of the case in light of all the evidence. A fact-finder model is typically a model of "the whole case". In criminal cases, fact-finder models serve to assess the probability of the prosecution hypothesis.

A Bayesian network can be more or less detailed. A more detailed network takes account of more circumstances and more dependencies between circumstances, and is therefore more complex. Selecting the level of detail is a trade-off between simplicity and nuance. This goes for expert models as well as fact-finder models. The distinction between expert models and fact-finder models is not a distinction about the level of detail. It is a distinction between two different purposes and two different questions.

This paper is concerned with fact-finder models, not forensic expert models. The pioneering work on fact-finder models was made by Norman Fenton, David Lagnado and Martin Neil (Fenton, Neil & Lagnado 2011, 2013a, 2013b, Fenton & Neil 2019; Lagnado 2021), and further contributions to fact-finder modeling in various situations have been made by Marcello Di Bello (2021), Christian Dahlman (2022, 2024)

and Hylke Jellema (2024).[1] An example of fact-finder modeling, taken from Fenton et al (2020), is displayed in Fig. 1. It is a model of the *Simonshaven Case*, a Dutch criminal case where a man called Ed was prosecuted for killing his wife, Jenny, in a forest near Rotterdam. Ed claimed that he was innocent and said that he and his wife had been assaulted by a stranger who suddenly came out of the bushes and attacked them as they were taking a walk. The Bayesian network in Fig. 1 models the probabilistic dependencies between the hypotheses at issue in the case, and their relation to various pieces of evidence, for example the connections between the hypotheses that Ed had a motive and owned a gun and the hypothesis that he killed his wife.

The *Simonshaven Case* is an example of a criminal trial with competing scenarios. The prosecution presented a scenario in which Ed killed his wife, ditched his gun at a pump station and invented a story where he and his wife had been attacked by a stranger coming out of the bushes. The defense suggested an alternative scenario: the killer was a madman who attacked people in recreational areas. The hypotheses in the prosecution-scenario can be found on the left side of the model in Fig. 1, and the defense-scenario, circumscribed and marked as "alternative narrative", is placed on the right side.

These two scenarios are implicitly present in the model. If one goes through the network node by node, following the direction of the arrows, parts of the scenarios pop up here and there. The arrows in a Bayesian network represent probabilistic dependencies, but they are conventionally modeled pointing in the direction of causality. The step from one node to a connected node in the direction of the arrow will, therefore, often correspond to a part of a story that serves as one of the scenarios in the case. For example, the sequence of hypotheses represented by the nodes "Defendant criminal background", "Defendant owned gun" and "His gun used in killing" reads as a part of the prosecution narrative. It should be noted, however, that the scenarios of the case are not explicit in the Bayesian network. As we have seen, the purpose of a Bayesian network is to model the probabilistic dependencies between the hypotheses at issue and the various pieces of evidence. The purpose is not to highlight scenarios, and the scenarios of the case are therefore often hidden as tacit background assumptions, buried in a forest of nodes and arrows. This is a fundamental difference between Bayesian networks and scenario analysis.

In the scenario approach to legal evidence, the scenarios of the case are explicitly modeled. A scenario is modeled as a series of events on a timeline, and alternative scenarios are contrasted with each other on the same timeline, showing where events diverge. The pioneering work in the scenario approach was made by Willem Wagenaar, Peter van Koppen, and Hans Crombag (Wagenaar et al 1993), and further contributions have been made by Floris Bex (2011, 2020), Bart Verheij (2016) and Anne Ruth Mackor (van Koppen & Mackor 2020; Mackor & van Koppen 2021). The scenario approach is not a formal theory. Scenario models are not constructed according to underlying mathematical definitions.

Figure 2 shows an example of scenario modeling in the *Simonshaven Case*, taken from Bex (2020). The prosecution-scenario is represented by a horizontal sequence, and the madman scenario is represented as an alternative sequence of events below.

[1] It has been debated among scholars to what extent Bayesian networks are practically useful in the context of legal evidence. For a discussion on the challenges, see Fenton & Lagnado (2021).

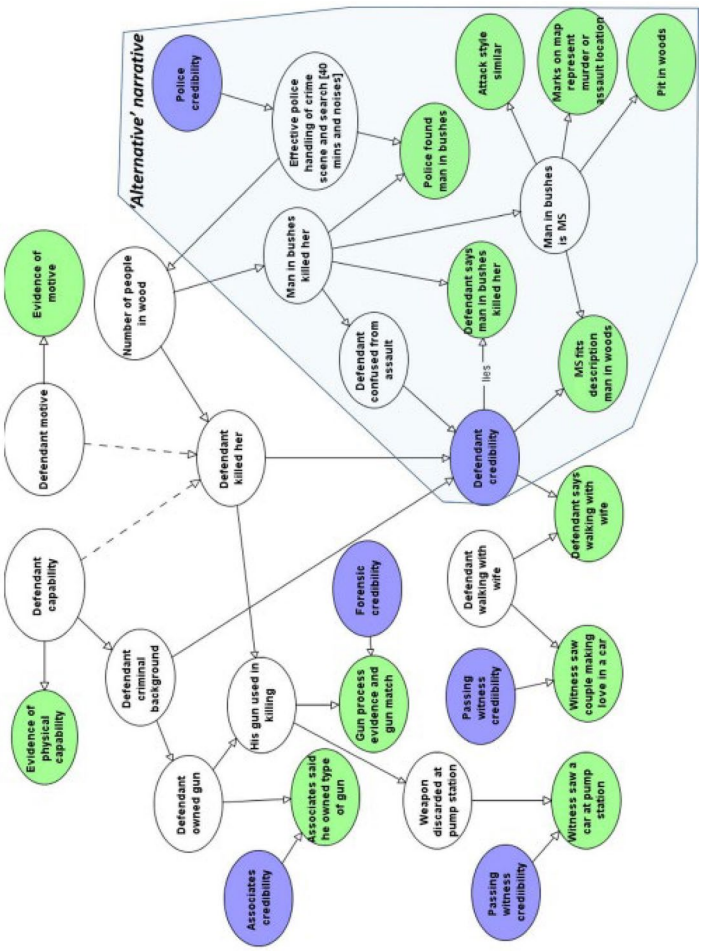


Fig. 1 Example of a Bayesian modeling in a legal case (*Simonshaven*)

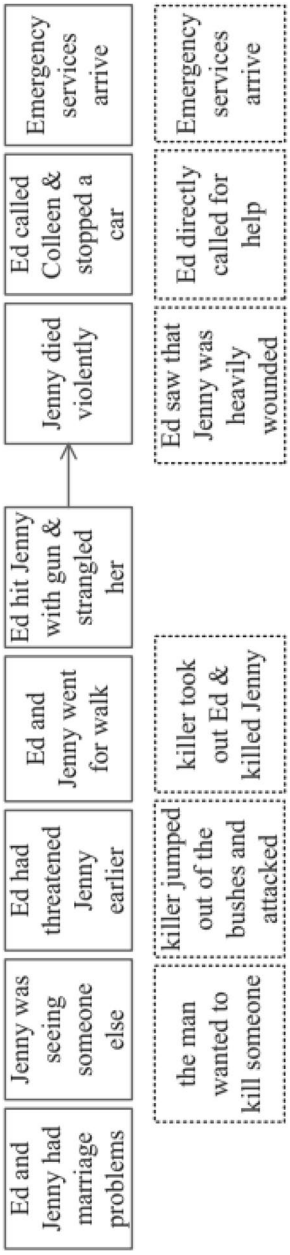


Fig. 2 Example of scenario modeling in a legal case (*Simonshaven*)

Scenario analysis is helpful in the assessment of legal evidence. Making alternative scenarios explicit helps the fact-finder to consider alternative explanations at every step of the assessment (Mackor & van Koppen 2021). As we have seen, Bayesian networks model alternative hypotheses but do not focus on scenarios. Bayesian modeling can therefore be supported by scenario analysis as it promotes the consideration of alternative hypotheses.

Scenario analysis is helpful to Bayesian modeling in several ways. For the purpose of this paper, we make a distinction between two kinds of assistance that we will refer to as *formal assistance* and *informal assistance*. Formal assistance occurs when a scenario model identifies a connection between scenario elements and this connection is then imported without loss into a Bayesian model as a connector between nodes in the network. It can either be a *probabilistic connection* or a *spatio-temporal constraint*. There is a probabilistic connection between events in a scenario when one event makes the other event more (or less) probable, for example when an event giving the defendant a motive to commit a certain act increases the probability of the defendant committing that act. There is a spatio-temporal constraint between scenario events when they are logically incompatible with each other, for example requiring the defendant to be in two different locations at the same time. Formal assistance is quantitative in the sense that the connections identified in the scenario model become numbers in the node probability tables of the Bayesian network. Probabilistic connections become probabilities between zero and one and spatio-temporal constraints become probabilities of zero for logically incompatible events.

Informal assistance occurs in various ways that are more qualitative in nature, for example as critical questions to keep in mind when modeling Bayesian networks, e.g. "what kind of evidence would we expect to see if the scenario is true?", "does the scenario have gaps in its timeline?", "has every alternative explanation been considered?", *etcetera* (van Koppen & Mackor 2020). In this paper, we will explore how scenario analysis can give formal assistance to Bayesian modeling and focus on spatio-temporal constraints. We do not claim that informal assistance is less important, but it is an issue for a different paper.

We will explore how scenario analysis can help Bayesian modeling to identify *spatio-temporal constraints* that need to be respected in Bayesian networks. As an example of a spatio-temporal constraint, consider a case where the following holds. If the defendant was at home at a certain time, T1, he did not have the opportunity to commit the crime at T2, since the number of minutes between T1 and T2 would not have been sufficient to get to the crime scene from his house. This spatio-temporal constraint can be modeled by setting P(defendant at crime scene at T2 = true|defendant at home at T1 = true) = 0 in the node probability table for "defendant at crime scene at T2" in Fig. 3.

Spatio-temporal constraints are often about *opportunity* but can also concern the *means* to commit the crime in question. For example, the defendant in a murder case would not have had the means to commit the crime if he did not have the murder weapon in his possession.

In the following sections, we will explore three different methods for constructing scenario-based Bayesian networks: scenario-nodes (Sect. 3), non-Boolean time-nodes with event-states (Sect. 4), and non-Boolean event-nodes with time-states

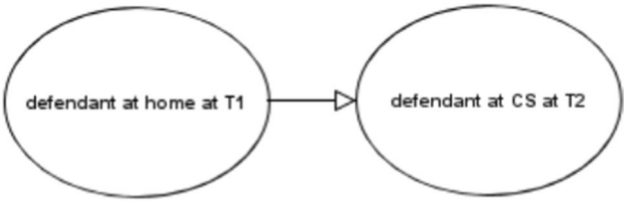


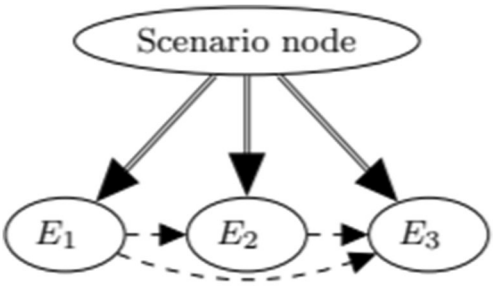
Fig. 3 Example of spatio-temporal constraint

(Sect. 5). As we mentioned above, scenario-nodes is a method that has been suggested and developed by other scholars, while non-Boolean time-nodes and event-nodes are innovations that have not previously been explored in the literature. For a formal mathematical definition of non-Boolean nodes, see the Appendix.

3 Scenario-nodes

Bayesian networks with scenario-nodes is a modeling method that has been proposed and developed by Charlotte Vlek, Henry Prakken, Silja Renooij and Bart Verheij (Vlek et al 2014, 2016). It takes a certain scenario that entails a number of elements (E_1, E_2 …) and turns it into a Bayesian network by modeling the scenario as a node and each element as a separate connected node. See Fig. 4, from Vlek et al 2016. As an example, the scenario-node could represent the following scenario: "the defendant stole a handbag from a parked car by smashing the car window and grabbing the bag", and the element-nodes could be the following events: "the defendant intended to take the handbag" (E_1), "the defendant smashed the car window" (E_2), and "the defendant took the handbag" (E_3). The dependencies are modelled with the scenario-node as a parent-node, i.e. the arrows go from the scenario-node to the elements. This means that the node probability table for each element is conditioned on the scenario-node being true or false. Since the elements are entailed in the scenario,

Fig. 4 Modeling with scenario-node



Scenario node	T	F
E_1 =T	1	…
E_1 =F	0	…

each element-node must be true if the scenario-node is true (see the node probability table in Fig. 4). However, it is not necessarily the case that an element is false if the scenario-node is false.

These dependencies between the scenario-node and the element-nodes lead to a propagation effect in the network: when one element of the scenario becomes more probable the probability of the other elements increases as well. Vlek and her colleagues refer to this as "the effect of coherence of a scenario" leading to "the transfer of evidential support" between the elements (Vlek et al 2016, 289).

The problem with this propagation effect is that it is not true for every relation between two elements in every scenario. Some elements are probabilistically dependent and should give rise to a propagation effect, for example if it becomes more probable that "the defendant intended to take the handbag" (E_1) it also becomes more probable that "the defendant smashed the car window" (E_2), but there are also elements in scenarios that are probabilistically independent. Consider the following example (*Hostage Case*). A terrorist group has taken a hostage. Every day the leader of the terrorist group flips a coin. If it comes up heads, a finger on the hostage is severed. If it comes up tails, no finger is severed on that day. Imagine that two days have passed and the police are contemplating the worst-case-scenario that the hostage has lost two fingers (the scenario in which the coin-flip on day one came up heads as well as the coin-flip on day two). The police learn that on the evening of day one a severed finger (left pinky) stemming from the hostage was found, indicating that the coin-flip on day one came up heads. What is the probability that the coin-flip on day two also came up heads, and that the hostage has lost another finger? The correct answer to this question is 50%, since the coin-flips are probabilistically independent, but a Bayesian network of the *Hostage Case* modeled with a scenario-node gives a different, incorrect answer. See Fig. 5. The scenario-node "day one: coin-flip heads finger severed, day two: coin-flip heads another finger severed" has two elements: "day one: coin-flip heads finger severed" and "day two: coin-flip heads finger severed", and the first element is supported by a piece of evidence: "day one: left pinky found severed". The left model in Fig. 5 shows the network before the evidence has been observed,[2] and the right model shows what happens when the police learn about the discovered pinky finger and the evidence-node is instantiated. The probability of "day one: coin flip heads finger severed" correctly goes to 100%, but the probability of "day two: coin flip heads finger severed" incorrectly increases from 1/2 to 2/3.

This is a mathematical flaw in the scenario-node approach. To get rid of this problem we need to get rid of the scenario-node and model the dependencies between elements in a scenario directly between the element-nodes, like the dashed arrows in Fig. 4. In the following sections, we will explore two methods for doing this.

[2] The priors in the element-nodes are bit tricky, since they are conditioned on their own conjunction with the other element. P(day one coin flip heads = true|day one coin flip heads and day two coin flip heads = false) = P(day two coin flip heads = true|day one coin flip heads and day two coin flip heads = false) = 1/3. This may appear counter-intuitive and resembles the so-called *Monty Hall Problem* (cf. Selvin 1975).

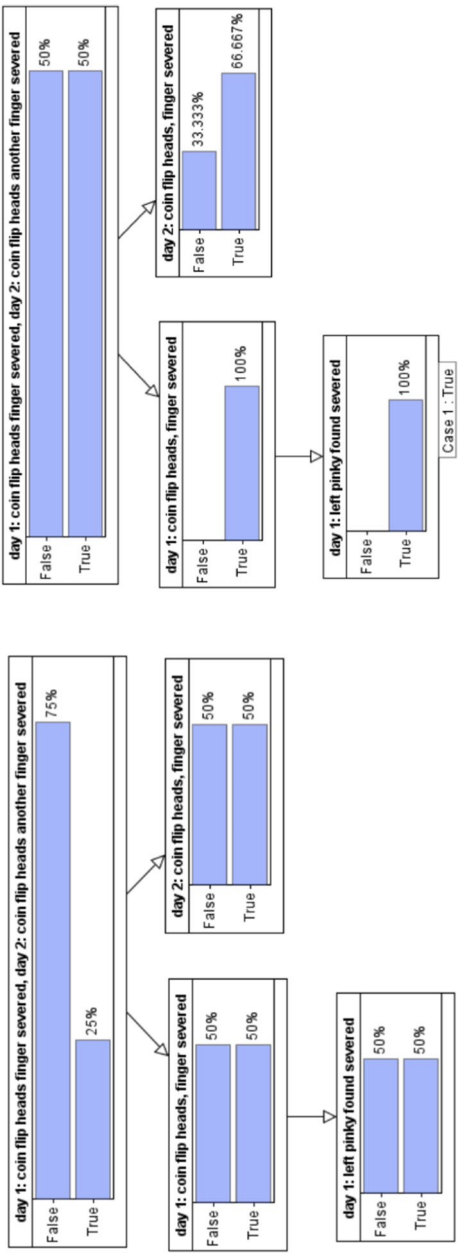


Fig. 5 Example incorrect propagation due to scenario-node (*Hostage Case*)

4 Time-nodes with event-states

To see how scenario models assist Bayesian modeling we need to start with scenario models with alternative scenarios as events on a timeline. The generic structure of a scenario model is displayed in Fig. 6. Scenarios (S1, S2 …) are structured in relation to a shared timeline, and each scenario has events (E) corresponding to times (T1, T2 …). For example, "E(S2)(T3)" refers to the event that happened according to scenario 2 at the time T3.

If an event in one scenario is incompatible with an event in the other scenario, this relation can be modelled as a *spatio-temporal constraint* between these events. As an example, displayed in Fig. 7, E(S2)(T2) and E(S2)(T3) could be incompatible with E(S1)(T4).

Such constraints ensue in the following example scenario (*Case of the Two Shootings*). On the same day, two shootings happen in the same city, one hour apart. The first shooting occurs at 12:00 in a parking lot, the second shooting happens at 13:00 in a park. CCTV footage from the parking lot shows two men firing at each other. At 12:00, one of them is hit and falls to the ground. He is later identified as Alex. The other man flees the scene, but his face is captured on camera. He is later identified as Carlos. Alex remains wounded on the ground. At 12:15, the CCTV footage shows another man arriving at scene, kneeling beside Alex. He takes the gun from Alex's hand, puts it in his pocket and leaves. His face is not captured on camera. At 12:20, an ambulance arrives and takes Alex to hospital. At 13:00, the second shooting happens in a park in a different part of the city. A man fires at another man, who falls dead to the ground, and the shooter leaves the scene. CCTV footage does not reveal the shooter's face but clearly shows that he is left-handed.

There are two connections between the shootings. The man shot dead in the park is identified as Carlos, and the bullets that Carlos was killed with come from the same gun as the bullets that were fired against Carlos an hour before by Alex in the parking lot. Since Alex was in the hospital at 13:00 when Carlos was killed in the park, the police assume that the man who kneeled beside Alex at 12:15 and took Alex's gun then shot Carlos in the park with it, as revenge for shooting Alex. The police suspect that the killer was Alex's brother Bob, who is left-handed. When Bob is questioned by the police, he claims that he is innocent. Bob says that he was at home at 12:00 when Alex was shot. He claims that he is not the man in the CCTV footage kneeling beside Alex at 12:15. Bob says that he was still at home at that time, and that his mother can verify this. Bob says that he went to the beach and was nowhere near the park when Carlos was killed. He says that he arrived at the beach at 12:45 and met his sister who can vouch for that. According to Bob's story, he was walking by himself on the beach at 13:00 when Carlos was killed. Figure 8 shows a scenario model of the case, with spatio-temporal constraints.

The dashed line between "Bob at home" and "Bob shoots Carlos in the park with Alex gun" represents the spatio-temporal constraint that if Bob was at home at 12:15, he could not have picked up Alex's gun in the parking lot, and could not have shot Carlos with it.[3] The dashed line between "Bob at the beach" and "Bob shoots Carlos

[3] Given the timeframe and the distance between the parking lot and Bob's house, it is not possible that someone else picked up the gun and gave it to Bob.

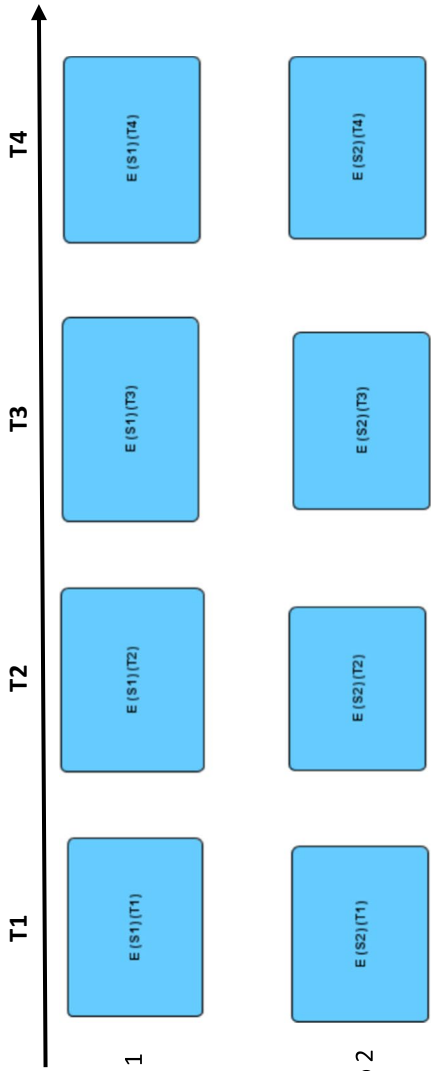


Fig. 6 Scenario model (generic)

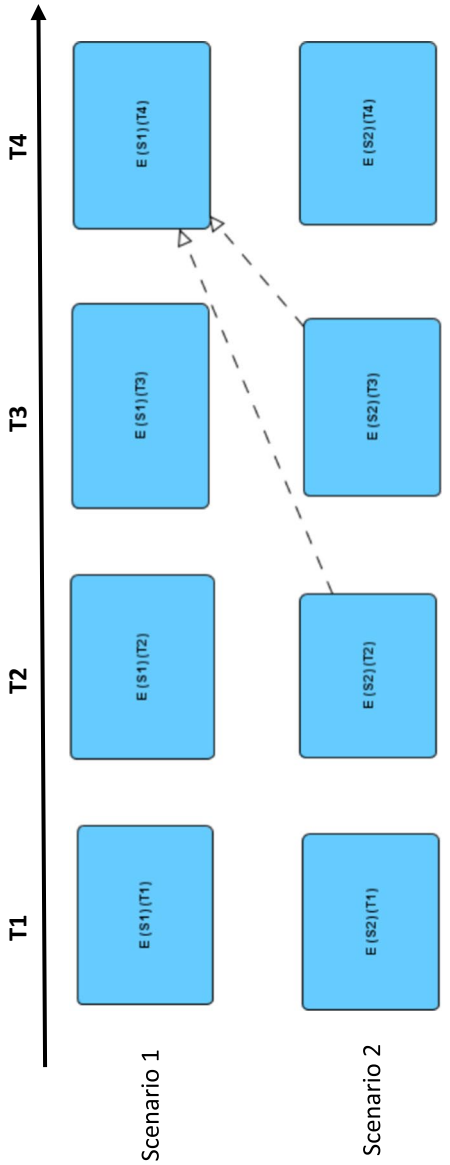


Fig. 7 Scenario model with spatio-temporal constraints

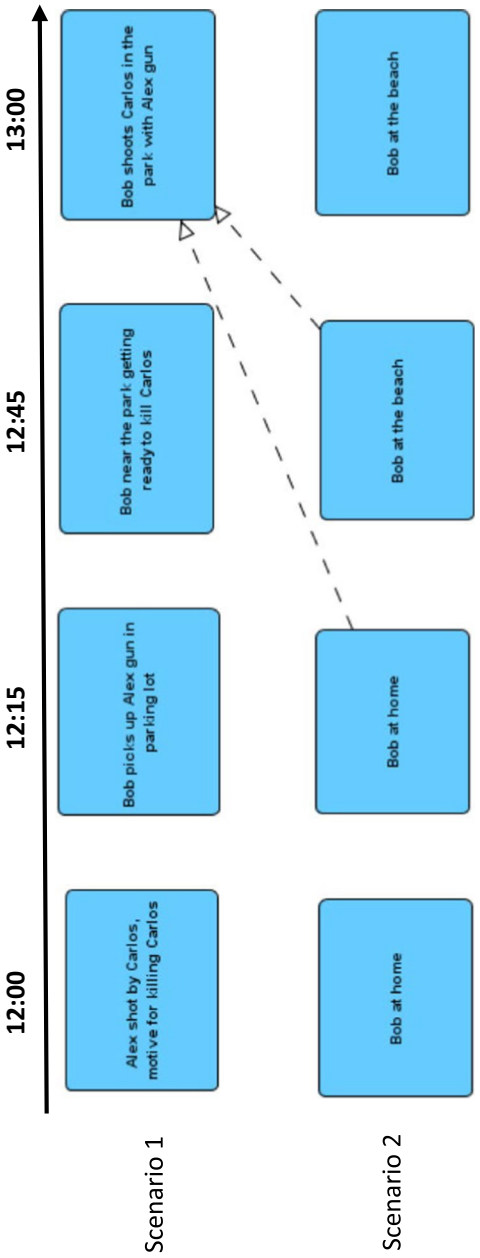


Fig. 8 Scenario model (*Case of the Two Shootings*)

in the park with Alex gun" represents the spatio-temporal constraint that if Bob was at the beach at 12:45, he could not have shot Carlos in the park at 13:00, since 15 min is not enough time to get from the beach to the park.

The evidence in a legal case consists of various pieces of evidence connected to the events in the scenarios. In the *Case of the Two Shootings*, the prosecution-scenario is supported by two pieces of evidence: the matching bullets and the CCTV-footage from the park showing that the killer was left-handed. The fact that the bullets from the two shootings match supports the hypothesis that Carlos was killed with Alex's gun, which further supports the hypothesis that the shooting of Alex was the motive for the shooting of Carlos, which supports that Bob shot Carlos, since Bob is Alex's brother. The killer being left-handed gives additional support to the hypothesis that it was Bob, since he is left-handed. How much support does this evidence give to the case for the prosecution? This is not easy to assess, especially with regard to the matching bullets, since the support that this evidence gives to the hypothesis that Bob killed Carlos goes through several intermediate hypotheses.

The defense-scenario is also supported by two pieces of evidence: the testimonies of Bob's mother and sister. How much support does this give to the defense-scenario, considering they could be lying to protect Bob? The fact-finder must consider all of the evidence on both sides, and assess how probable it is, all things considered, that Bob killed Carlos.

The scenario model does not offer assistance in this endeavor. It does not provide any formal tools for probability assessments. The formal assistance for this task is provided by Bayesian modeling. It is therefore very helpful to combine scenario modeling with Bayesian modeling in a scenario-based Bayesian network that models spatio-temporal constraints together with probabilistic dependencies. This can be done with *non-Boolean time-nodes with event-states*. Figure 9 shows how the scenarios in the *Case of the Two Shootings* can transformed into non-Boolean time-nodes for a Bayesian network.

The difference between Boolean and non-Boolean nodes is that Boolean nodes have two possible states – true or false – while non-Boolean nodes can have as many different states as needed for the modeling of the case. A non-Boolean time-node represents a time on the timeline and its states are alternative events that took place at that time according to the alternative scenario of the case. For the modeling of the *Case of the Two Shootings*, the nodes represent the times 12:00, 12:15, 12:45 and 13:00, and the states represent the alternative events that occurred at these times according to the two scenarios. Non-Boolean time-nodes are an elegant way to handle spatio-temporal constraints in Bayesian networks. They make the scenarios visible and require fewer nodes than traditional Boolean modeling.

Events in alternative scenarios are sometimes defined in a way that exhausts all possibilities. This is, for example, the case with the claim in the prosecution-scenario that Bob killed Carlos and Bob's claim that he did not. The node representing what happened at 13:00 is therefore only equipped with two states. The man that shot Carlos was either Bob or someone else. There is no room for a third possibility. However, events in alternative scenarios are often non-exhaustive, leaving room for other possibilities. For example, at 12:45, Bob being near the park (prosecution-scenario) or Bob being at the beach (defense-scenario) does not exhaust all possibilities. It is

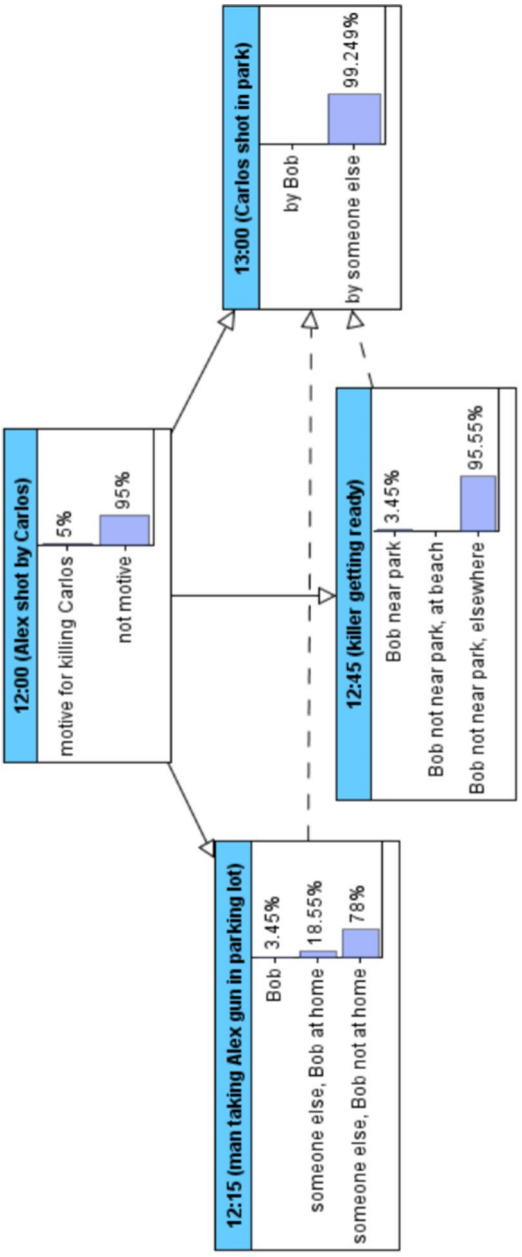


Fig. 9 Non-Boolean time-nodes with event-states (*Case of the Two Shootings*)

possible that Bob was at some other location that is not near the park, and that the scenarios put forward by the prosecution and defense are both wrong. The time-node for 12:45 is therefore equipped with three states.

In Fig. 9, spatio-temporal constraints are represented with dashed arrows and probabilistic dependencies with solid arrows. The spatio-temporal constraints become zeros in the node probability tables. For example, in the node probability table for 13:00, P(Carlos shot by Bob = true|Bob not near park, at beach = true) = 0. The arrows from 12:00 to the other nodes are solid since the relation is probabilistic: that the shooting of Alex was the motive for the shooting of Carlos increases the probability that Bob took Alex's gun and went to the park and killed Carlos, but it does not make it impossible that someone else took the gun and killed Carlos. In Fig. 10, the evidence and the intermediate hypothesis have been added to the model. Green nodes are prosecution-evidence and red nodes are defense-evidence.

A Bayesian network assists in the assessment of the evidence by helping to identify the probabilistic dependencies in the case and test how different assumptions affect the posterior probability. Given the assumptions that have been put into the node probability tables in Fig. 10, the probability that Bob shot Carlos is approximately 60%. These assumptions are of course debatable. For example, it is not easy to say how likely it is that Bob's mother would lie in court and say that he was at home at 12:15, if this is not true. The model in Fig. 10 assumes that there is a 20% probability that she would do this. A fact-finder who thinks that the probability is actually higher (or lower) can adjust the value in the node probability table and see how this affects the probability that Bob was at home at 12:15 and the probability that he shot Carlos in the park at 13:00. The Bayesian model helps the fact-finder to identify what assumptions are necessary to bring the posterior probability to the degree needed for a conviction, so that the fact-finder can make an informed decision on whether the assumptions needed are realistic and reasonable.

5 Event-nodes with time-states

In some legal cases, modeling alternative scenarios is not a question of whether one event or another occurred at a certain time, but whether an event happened at one time or another. For example, it could be certain that the defendant was at a restaurant on the night when the crime was committed, but uncertain at what time the defendant left the restaurant. Depending on the time, the defendant either had opportunity to commit the crime, or did not have opportunity. Such spatio-temporal constraints can also be modeled with alternative scenarios on a time-line.

As an illustration, consider the following example (*Burglary Case*), where the possible scenarios are modeled in Fig. 11. A burglary is committed in an office building, at some time around 21:00 or 22:00 in the evening when the building is not in use. The police suspect that a woman called Betty was the burglar. Betty claims that she is innocent. She says that she had dinner by herself at a restaurant and then went to a night club. The police speak to a waiter at the restaurant, who confirms that Betty was there on the night in question and left at 21:00. The police also speak to a bartender at the night club, who confirms that Betty was there, and came in at 21:30. Given the locations of the

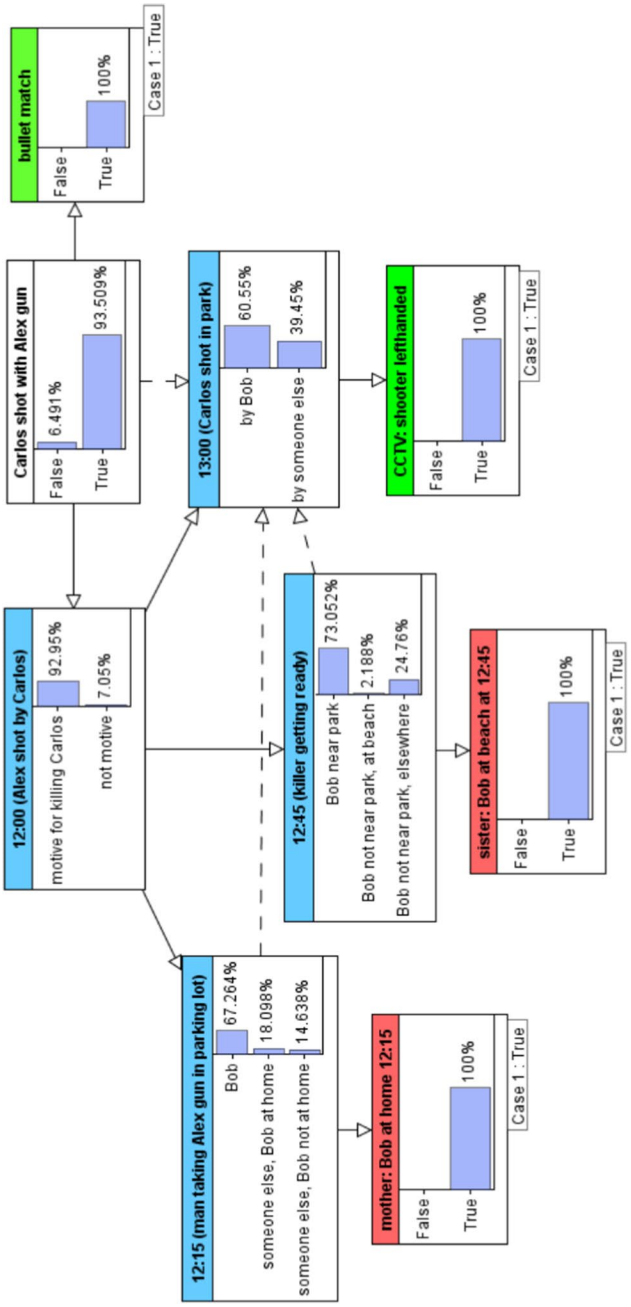


Fig. 10 Bayesian network with time-nodes and evidence (*Case of the Two Shootings*)

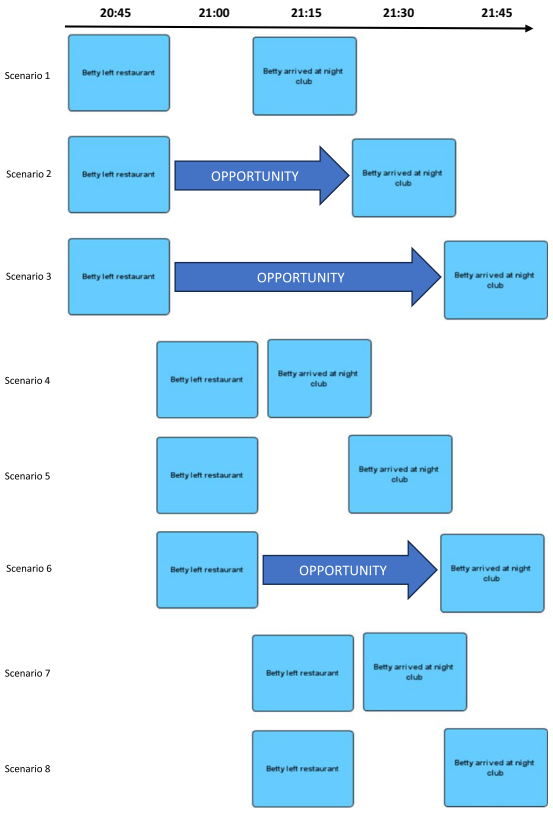


Fig. 11 Scenario model (*Burglary Case*)

restaurant, the office building and the night club, it would have taken Betty more than 30 min to commit the burglary between her visit at the restaurant and her visit at the night club, so she did not have the opportunity if the waiter and the bartender remember the time correctly. However, 45 min would have been enough time, so if Betty, for example, left the restaurant 15 min earlier, or arrived at the night club 15 min later, she did have opportunity. The possibility that the waiter and/or the bartender is wrong about the time must therefore be considered. The possible scenarios with 15-min-deviations in each direction are modeled in Fig. 11. As the model shows, Betty had the opportunity to commit the burglary in three of these scenarios (scenario 2, 3 and 6).

The scenario model is helpful for identifying the spatio-temporal constraints related to opportunity, but it does not provide the formal tools for assessing the probability that Betty had opportunity, given the testimonies of the waiter and the bartender. As we have seen, the formal assistance for this assessment is provided by Bayesian modeling.

A Bayesian model for the *Burglary Case* can be constructed with time-nodes, but a simpler way to model the spatio-temporal constraints in the *Burglary Case* is to employ event-nodes with time-states. Such event-nodes are like time-nodes turned on their head. They represent events, and the states of each node represent alternative times when that event took place, according to the alternative scenarios of the case. Figure 12 shows how the *Burglary Case* can be modeled in this way.

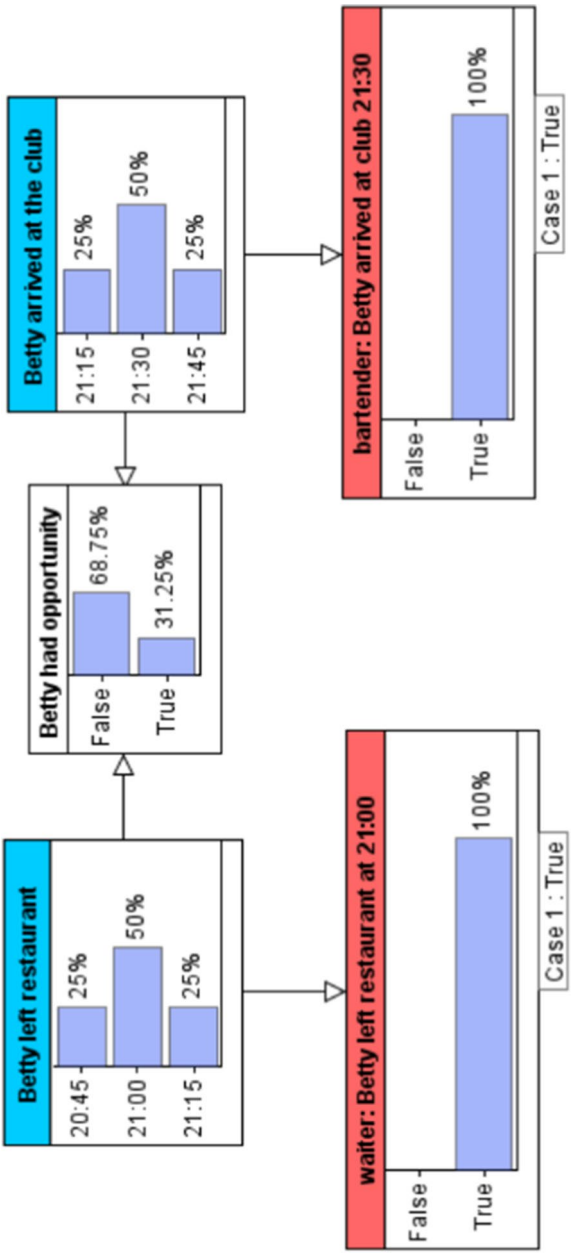


Fig. 12 Bayesian network using event-nodes with time-states

In this model, it is assumed that there is a 25% probability that a witness like the waiter or the bartender would be 15 min wrong about the time. Given this assumption, the probability that Betty had opportunity to commit the burglary is approximately 31%. As we have seen above, the Bayesian model helps the fact-finder to identify the probabilistic dependencies in the case, and how certain assumptions affect the posterior probability.

6 Conclusions

In this paper we have explored how the construction of Bayesian networks for legal evidence can be assisted by scenario models. We have looked at *fact-finder models*, where "the whole case" is modeled from the perspective of the fact-finder, and we have studied *formal assistance*, where a connection between elements in a scenario model is imported into a Bayesian model as a connection between nodes. We have focused, specifically, on *spatio-temporal constraints* in scenario-based Bayesian networks.

Our investigation makes several theoretical contributions to the literature on the modeling of legal evidence. It has been observed before that scenario thinking can be of informal assistance in the construction of Bayesian networks, but our investigation is the first to demonstrate exactly how scenario models can be of formal assistance. We have shown how spatio-temporal constraints can be identified in scenario models and transferred into Bayesian models of legal cases.

Our paper makes a novel contribution to the literature on fact-finder models. We have introduced two new methods for constructing scenario-based Bayesian networks: time-nodes with event-states and event-nodes with time-states.

A *non-Boolean time-node with event-states* represents a time on the timeline and its states are alternative events that took place at that time according to the alternative scenarios of the case.

A *non-Boolean event-node with time-states* represents an event and its states represent alternative times when that event took place, according to the alternative scenarios of the case.

These modeling methods handle spatio-temporal constraints in a clear and simple way. In contrast to traditional Bayesian fact-finder models, the scenarios become visible in the model, which makes it easier to see what is going on.

It comes naturally to legal fact-finders to think in terms of scenarios, but they struggle when it comes to Bayesian networks. The modeling methods that we have introduced in this paper should make it easier for fact-finders to transform scenarios into Bayesian networks.

Appendix

Formal mathematical definitions of Bayesian networks

1. Graphical Structure

A Bayesian network is defined over a directed acyclic graph (DAG) G=(V, E), where V={X_1, X_2, …, X_n} is a set of random variables and E is a set of directed edges. The graph contains no directed cycles.

For each variable Xi, define its parent set as:

$$\mathrm{Pa}\left(\mathrm{X_i}\right) = \left\{\mathrm{X_j inV} \middle| \left(\mathrm{X_j} -> \mathrm{X_i}\right) \mathrm{inE}\right\}.$$

2. Probabilistic Semantics

The joint probability distribution over V factorizes according to the graph structure:

$$\mathrm{Pa}\left(\mathrm{X_1}, \mathrm{X_2}, \cdots, \mathrm{X_n}\right) = \text{product of i over } \mathrm{P}\left(\mathrm{X_i} \middle| \mathrm{Pa}\left(\mathrm{X_i}\right)\right).$$

3. Local Markov Property

Each variable X_i is conditionally independent of its non-descendants given its parents:

$$\mathrm{X_i} \perp \mathrm{NonDescendants}\left(\mathrm{X_i}\right) \middle| \mathrm{Pa}\left(\mathrm{X_i}\right).$$

4. Global Markov Property

For disjoint subsets A, B, C of V, if A and B are d-separated by C in G, then:

$$\mathrm{A} \perp \mathrm{B} \middle| \mathrm{C}.$$

5. Parameterization

Each variable X_i is associated with a conditional probability distribution $P(X_i|Pa(X_i))$. For discrete variables, this is represented by a conditional probability table (CPT).

6. Complete Definition

A Bayesian network is a pair B=(G, Θ), where G is a DAG and Θ is a set of parameters defining the conditional distributions $P(X_i|Pa(X_i))$, such that:

$$\mathrm{P}\left(\mathrm{V}\right) = \text{product of i over } \mathrm{P}\left(\left(\mathrm{Xi}|\mathrm{Pa}(\mathrm{Xi})\right)\right).$$

7. Use of Bayes' Theorem in Bayesian Networks

Bayes' theorem provides the fundamental rule for updating probabilities in light of new evidence. For events A and B with P(B)>0, Bayes' theorem states:

$$\mathrm{P}\left(\mathrm{A} \middle| \mathrm{B}\right) = \left[\mathrm{P}\left(\mathrm{A} \middle| \mathrm{B}\right) \mathrm{P}\left(\mathrm{A}\right)\right] / \mathrm{P}\left(\mathrm{B}\right).$$

In a Bayesian network, this theorem is used to calculate the posterior probability of a hypothesis (H) given observed evidence (E). The posterior distribution is given by:

$$\mathrm{P}\left(\mathrm{H}|\,\mathrm{E}\right) = \mathrm{P}\left(\mathrm{H},\mathrm{E}\right)/\mathrm{P}\left(\mathrm{E}\right).$$

Using the factorization property of the Bayesian network, the joint distribution P(H,E) can be expressed as a product of local conditional probabilities:

$$\mathrm{P}\left(\mathrm{H},\mathrm{E}\right) = \text{product of i over } \mathrm{P}\left(\mathrm{H_i}|\mathrm{Pa}\left(\mathrm{H_i}\right)\right).$$

This process underlies all inference algorithms in Bayesian networks.

8. Non-Boolean Nodes

The random variables X_i in a Bayesian network are not restricted to Boolean domains. Each variable X_i takes values in a measurable space (Ω_i, F_i).

For discrete non-Boolean variables, X_i may take values in a finite or countable set $\{x_{i1}, x_{i2}, \ldots, x_{ik}\}$, and its conditional distribution is given by a conditional probability table:

$$\mathrm{P}\left(\mathrm{X_i} = \mathrm{x_{ij}}|\,\mathrm{Pa}\left(\mathrm{X_i}\right)\right).$$

Authors' contribution The authors have contributed equally to this paper.

Funding Open access funding provided by Lund University. Research conducted in the research project "Preventing Miscarriages of Justice" funded by the Dutch Research Council (NWO), grant nr 406.21.RB.004.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

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

Bex F (2020) The hybrid theory of stories and arguments applied to the Simonshaven Case. Top Cogn Sci. https://doi.org/10.1111/tops.12426

Dahlman C, Kolflaath E (2022) Causal models versus reason models in Bayesian networks for legal evidence. Synthese. https://doi.org/10.1007/s11229-022-03953-y
Dahlman C (2024) The Virtue of Being Disagreeable. In: Feteris E et al (eds) Legal Argumentation – Reasoned Dissensus and Common Ground. Boom, Rotterdam
Di Bello M (2021) A probabilistic analysis of cross-examination using Bayesian networks. Philos Issues. https://doi.org/10.1111/phis.12209
Fenton N, Neil M (2019) Risk assessment and decision analysis with Bayesian networks, 2nd ed. CRC Press
Fenton N, Neil M, Lagnado D (2013a) A general structure for legal arguments about evidence using Bayesian networks. Cogn Sci. https://doi.org/10.1111/cogs.12004
Fenton N, Neil M, Lagnado D (2013b) Legal idioms – A framework for evidential reasoning. Argument Comput. https://doi.org/10.1080/19462166.2012.682656
Fenton N, Neil M, Yet B, Lagnado D (2020) Analyzing the Simonshaven Case using Bayesian networks. Top Cogn Sci. https://doi.org/10.1111/tops.12417
Fenton N, Lagnado D (2021) Bayesianism – objections and rebuttals. In: Dahlman C (ed) Philosophical foundations of evidence law. Oxford University Press
Fenton N, Neil M (2011) Avoiding Legal Fallacies in Practice Using Bayesian Networks. Australian J Legal Philos 36
Jellema H (2024) Perpetrator knowledge - a Bayesian account. Law Probab Risk. https://doi.org/10.1093/lpr/mgae009
Lagnado D (2021) Explaining the evidence. Cambridge University Press
Mackor AR, van Koppen P (2021) The scenario theory about evidence in criminal law. In: Dahlman C (ed) Philosophical foundations of evidence law. Oxford University Press
Selvin S (1975) A Problem in Probability, The American Statistician 29
Taroni F, Biedermann A, Bozza S, Garbolino P, Aitken C (2014) Bayesian networks for probabilistic inference and decision analysis in forensic science, 2nd ed. Wiley
van Koppen P, Mackor AR (2020) A scenario approach to the Simonshaven case. Top Cogn Sci. https://doi.org/10.1111/tops.12429
Verheij B, Bex F, Timmer S, Vlek C, Meyer J, Renooij S, Prakken H (2016) Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning. Law Probab Risk. https://doi.org/10.1093/lpr/mgv013
Vlek C, Prakken H, Renooij S, Verheij B (2014) Building Bayesian networks for legal evidence with narratives – a case study evaluation. Artif Intell Law. https://doi.org/10.1007/s10506-014-9161-7
Vlek C, Prakken H, Renooij S, Verheij B (2016) A method for explaining Bayesian networks for legal evidence with scenarios. Artif Intell Law. https://doi.org/10.1007/s10506-016-9183-4
Wagenaar W, van Koppen P and Crombag H (1993) Anchored Narratives – The Psychology of Criminal Evidence, St Martin's Press