



Causal models versus reason models in Bayesian networks for legal evidence

Christian Dahlman¹ · Eivind Kolflaath²

Received: 8 October 2021 / Accepted: 22 October 2022 / Published online: 14 November 2022

© The Author(s) 2022

Abstract

In this paper we compare causal models with reason models in the construction of Bayesian networks for legal evidence. In causal models, arrows in the network are drawn from causes to effects. In a reason model, the arrows are instead drawn towards the evidence, from *factum probandum* to *factum probans*. We explore the differences between causal models and reason models and observe several distinct advantages with reason models. Reason models are better aligned with the philosophy of Bayesian inference, as they model reasons for up-dating beliefs. Reason models are better suited for measuring the combined support of the evidence, and a prior probability of guilt that reflects the number of possible perpetrators is accommodated more easily with reason models.

Keywords Evidence · Bayesian network · Law · Causality · Reason · Prior probability

1 Introduction

There is an interesting mismatch between the philosophy of Bayesian inference and the development of Bayesian networks that has not received sufficient attention.¹ Bayesian philosophy is grounded in the subjective theory that probabilities are degrees of belief and represent Bayesian inferences as reasons for updating our degree of belief on the basis of new evidence. Yet, the development of Bayesian networks has not been aimed at modeling reasons for beliefs. The theory of Bayesian networks developed

¹ Thanks to David Lagnado for helpful discussions and comments on previous versions of this paper.

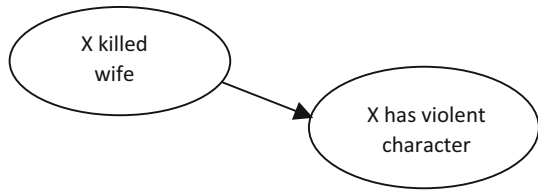
T.C. : Evidence in Law and Ethics.

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

¹ Faculty of Law, Lund University, Lund, Sweden

² Faculty of Law and Department of Philosophy, University of Bergen, Bergen, Norway

Fig. 1 ‘X has violent character’ modeled as reason for ‘X killed wife’



by Judea Pearl and his followers are causal models, where arrows between nodes represent causal relations (Fenton & Neil, 2018; Fenton et al., 2013; Lagnado, 2021; Pearl, 1988, 2009; Schweizer, 2015; Taroni et al., 2006; Vlek et al., 2016). There is a fundamental difference between reasons and causes. Reasons exist in our minds and serve as justifications for our beliefs. Causes exist in the external world and are not justifications.

Reasons can be modeled in graphs with nodes and arrows, just like causes. Such models have been developed in argumentation theory (see, for example, Walton et al., 2008), and Bayes’ rule can be applied to such models, to quantify reasons in probabilistic terms. So Bayesian networks can be used to model reasons instead of causal relations. Instead of drawing arrows to represent relations of cause and effect, the arrows can be drawn to represent the relation between *factum probandum* (needing reason) and *factum probans* (providing reason).

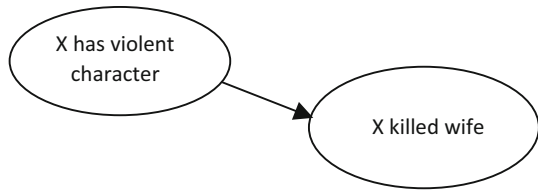
Here is a simple example of Bayesian reasoning with legal evidence. X’s wife has been found murdered. Before looking into the case, I believe that there is a 50% probability that X is the perpetrator, since my background information about the world tells me that 50% of murdered women are killed by their partner. Then I learn that X has a violent character, and up-date my degree of belief using Bayes theorem. I assume that $P(X \text{ has violent character} | X \text{ killed wife}) \approx 90\%$ and $P(X \text{ has violent character} | X \text{ did not kill wife}) \approx 5\%$, and therefore adjust the probability that X did killed his wife to 95% ($0.5/0.5 \times 0.90/0.05 \approx 0.95/0.05$). This Bayesian inference can be modeled in a simple Bayesian network (Fig. 1).²

In this model, the prior probability that X killed his wife before we learn about his violent character (50%) is inserted in the node probability table of the node ‘X killed wife’, and the conditional probabilities $P(X \text{ has violent character} | X \text{ killed wife})$ and $P(X \text{ has violent character} | X \text{ did not kill wife})$ go into the node probability table for ‘X has violent character’. Our reasoning when we learn about the violent character and up-date our degree of belief in the hypothesis that X killed his wife is modeled when we instantiate ‘X has violent character’ as true, and the probability of ‘X killed wife’ is up-dated according to Bayes theorem.

In this Bayesian network, the arrow goes from ‘X killed wife’ to ‘X has violent character’, since it models how ‘X has violent character’ is a reason for believing (to a certain degree) that ‘X killed wife’. If we, instead, would make a causal model, we would draw the arrow in the other direction, from ‘X has violent character’ to ‘X killed wife’ (Fig. 2).

² For an excellent introduction to Bayesian networks, see Fenton & Neil (2018).

Fig. 2 ‘X has violent character’ modeled as cause to ‘X killed wife’



In a causal model, arrows are drawn from cause to effect. As a consequence, the arrows must reflect the temporal order of the events. In our example, chronology dictates that the arrow must run from ‘X has violent character’ to ‘X killed wife’. Since X had a violent character before his wife was killed, it cannot be the case that the killing caused the violent character.

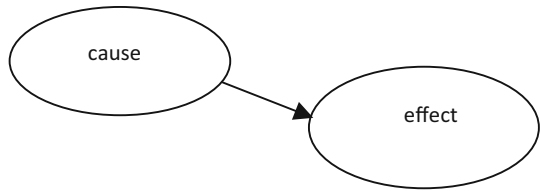
In the following, we will refer to these two ways of building Bayesian networks as *causal models* and *reason models*. In causal models the arrows between nodes are drawn in the direction of causality. In reason models, the arrows are drawn from *factum probandum* to *factum probans*. Quite often, the arrow will point in the same direction in both models, but there are situations, such as our example with X’s violent character, where an arrow will point in one direction in a causal model, while an arrow between the same nodes in a reason model will point in the other direction.

In the terminology for Bayesian networks, the node from which an arrow originates is a ‘parent’ to the node at which the arrow points, and the latter node is a ‘child’ to the former. In causal models, cause is modeled as parent, and effect as child. In reason models, *factum probandum* is modeled as parent, and *factum probans* as child.

In the existing literature on Bayesian network, causal models are completely dominant. They are so dominant that there is no discussion at all about ‘causal models’ versus ‘reason models’. We introduce this distinction in this paper, in the hope of opening up such a conversation. The purpose of this paper is to acknowledge the possibility to build Bayesian networks as reason models, clarify how they differ from causal models, and explore if they have some distinct advantages over causal models. We expect scholars who are used to causal models to be skeptical towards the use of Bayesian networks to build reason models, but we ask the reader to keep an open mind.

Readers should also keep in mind that causality plays no role in the formal definitions of the elements in a Bayesian network. When discussing what the arrows represent, we must distinguish between the formal (or purely probabilistic) meaning on the one hand and various informal, intended interpretations on the other. From a strictly formal point of view, the only meaning of the arrow in a Bayesian network is that it determines which of the two connected nodes is conditioned on the other.³

³ An arrow from A to B means that the relevant probabilities in the node probability table for node A are $P(A)$ and $P(\text{not-}A)$, while the relevant probabilities in the table for node B are $P(B|A)$, $P(B|\text{not-}A)$, $P(\text{not-}B|A)$ and $P(\text{not-}B|\text{not-}A)$. If we change the direction of the arrow, everything is reversed—the relevant probabilities in the tables are now $P(B)$ and $(\text{not-}B)$ for node B, and $P(A|B)$, $P(A|\text{not-}B)$, $P(\text{not-}A|B)$ and $P(\text{not-}A|\text{not-}B)$ for node A. For simplicity’s sake we will talk as if nodes themselves are probabilistically conditioned on other nodes, instead of mentioning the node’s probability table every time. What we mean by saying that ‘node A is conditioned on node B’ is that the content of node A is probabilistically conditioned

Fig. 3 Arrow from cause to effect

There is nothing in the formal mechanisms in Bayesian networks which forces us to interpret the arrows as anything beyond pure probabilistic dependency.

The paper is structured as follows. We start (Sect. 2) with a general discussion about the difference between causal models and reason models. Then (Sects. 3 to 5) we limit our discussion to legal evidence. In Sect. 3 we clarify what differences the choice between causal models and reason models makes in Bayesian networks of legal evidence. Section 4 is dedicated to the well-known ‘problem of the prior’, and we investigate how the presumption of innocence can be accommodated in causal models compared to reason models. The purpose of this section is to explore possible advantages and disadvantages with causal models compared to reason models.

2 Causal models versus reason models

In causal models, the probability relation between two nodes is modeled with the cause as a parent and the effect as a child. The direction of the arrow is the direction of causality, see Fig. 3.

This way of modeling is inherited from the pioneering work by Judea Pearl on Bayesian networks (Pearl, 1988, pp. 125–126). According to Pearl, ‘[t]he essential requirement for soundness and completeness is that the network be constructed *causally*’ (Pearl, 1988, p. 14).

This sounds rather strict, but causal modeling employs a notion of causality that is broader than our everyday use of the term (Lagnado, 2021, p. 39). Pearl defines causality as a form of listening: ‘X is a cause of Y if Y listens to X and decides its value in response to what it hears’ (Pearl et al., 2016, pp. 5–6). In our view, this definition is a bit too vague and metaphorical, but we will put that issue aside.

It should also be pointed out that causal modeling does not dictate that all arrows represent causal relations. Causal models allow for non-causal connections, for example arrows to so-called ‘constraint nodes’ (Fenton & Neil, 2018). Causal modeling only dictates that if there is a causal relation between two nodes, they should be connected with an arrow from cause to effect.

In reason models, by contrast, the direction of the arrows is determined by how propositions are related to each other as *factum probandum* (i.e., a proposition for which there are presented reasons) and *factum probans* (i.e., some reason for the *factum probandum*), see Fig. 4.

Footnote 3 continued

on the content of node B in the node probability table of the A-node. Since Bayesian networks are directed acyclic graphs, A being conditioned on B excludes the possibility that B is conditioned on A.

Fig. 4 Arrow from *factum probandum* to *factum probans*

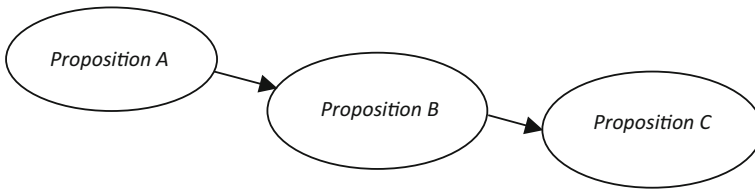
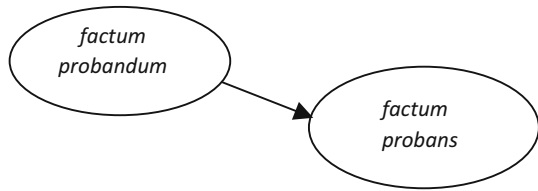


Fig. 5 Proposition with dual role (*factum probandum* and *factum probans*)

As we have seen, changing the direction of arrows changes what is conditioned on what. The fact that probability relations in a reason model may differ from those in a causal model gives rise to questions as to which probability relations are the relevant ones in the actual context. Such questions need to be considered carefully. Yet, the fundamental formal logic that guides reason modeling is the same as in causal modeling: Bayesian inference. Reason modeling follows Bayes’ Rule⁴ where *factum probandum* is the hypothesis (H) and *factum probans* is the evidence (E) applied as a reason to up-date the degree of belief in the hypothesis. In reason models, arrows from a node A point at reasons, e.g. B and C, for the truth or falsity of A, regardless of whether A is a possible cause of B and C. Thus, a *factum probandum* is represented by a parent node whose outgoing arrows point at some *factum probans* for the actual *factual probandum*. We mention also that like the concepts of cause and effect, we use the terms ‘factum probandum’ and ‘factum probans’ in a relational or relative sense. Just as an event B which is an effect of an event A may also be the cause of some event C, it is possible for a proposition to have two roles at the same time (Fig. 5).

Here, the proposition B plays the dual role of both being a *factum probans* in relation to A and a *factum probandum* in relation to C.

A fundamental difference between causal models and reason models is that a causal model is intended to represent the world itself, while reason models are representations of reasoning about the world. In causal models, nodes represent events in the world and causal relationships between them are represented by arrows pointing from cause to effect. In reason models, nodes represent propositional content, while arrows typically represent inferential relationships between propositions by pointing to a reason for updating the degree of belief in the other proposition.

Reason models also differ from causal models with respect to which nodes are candidates to be instantiated as true or false. In practice, which nodes are instantiated and which are not will depend on what is known, and which question one is trying to

⁴ $P(H)/P(-H) \times P(E|H)/P(E|-H) = P(H|E)/P(-H|E)$ (odds version).

answer, in each particular case. Sometimes we know that some event A has occurred and wonder whether it has caused some other event B, while other times we know that some event B has occurred and wonder whether it is some particular event A that caused it. In both cases, A and B are modelled as parent and child, respectively, but while the former type of situation implies that A will be instantiated, it is B which will be instantiated in the latter. In causal models, it therefore varies from case to case whether it is the parent node A or the child node B (if any) that is instantiated.

Not so with reason models. Here, child nodes always represent *facta probantia* (evidence in a broad sense) used in order to assess the probability of the *factum probandum* represented by the parent node. Reason models reflect that we make inferences from propositions we take to be true to propositions that require proof. The childless nodes, unlike all the other nodes, represent that which is regarded as true and not just probable. Accordingly, in reason models only childless nodes may be instantiated. Such a constraint does not apply to causal models, as we observed above. Moreover, in a reason model *every* childless node will eventually be instantiated, which is not necessarily the case with a causal model.

3 Modeling legal evidence

With this said about the differences between causal models and reason models in general, we now turn to the specific context of legal evidence. In Bayesian modeling of criminal evidence, the hypothesis-at-issue is the act that the defendant has committed according to the criminal charge put forward by the prosecution. As a shorthand, we will refer to this hypothesis simply as the ‘guilt-hypothesis’. With causal modeling an event occurring before another event in time will always be a parent, since causality cannot go backwards in time. For Bayesian models of criminal evidence, this means that e.g. character evidence and motive will be modeled as parents to the guilt hypothesis. Eyewitness testimony of the act stated in the guilt hypothesis will, on the other hand, be modeled as a child to the guilt hypothesis, and the same goes for forensic trace evidence (DNA, fingerprints etc.) of the act stated in the guilt hypothesis (Fig. 6).

The simplified structure in Fig. 6 serves to illustrate that character evidence and motive are modeled differently than eyewitness testimony and trace evidence but is too simplistic to be useful in an actual case. For instance, the overly simplified model in Fig. 6 does not provide any real help in assessing the probability that the defendant is guilty when we observe the character evidence and the motive. The model helps us to calculate how the probability that the defendant is guilty should be updated when we observe the eyewitness evidence and the DNA evidence, but when the nodes for character evidence and motive are instantiated as true the probability that the defendant is guilty will simply take the value of $P(\text{defendant guilty} | \text{character evidence defendant} \ \& \ \text{evidence of motive defendant})$ in the node probability table of the defendant-guilty-node as a prior. In other words, instead of helping us to assess the probability that the defendant is guilty given character evidence and motive, the network in Fig. 6 will require us to insert this probability in the node probability table, and then hand the same value back to us, when the nodes for character evidence and motive are instantiated as true. This is not the case with the more developed causal model in Fig. 7. Here,

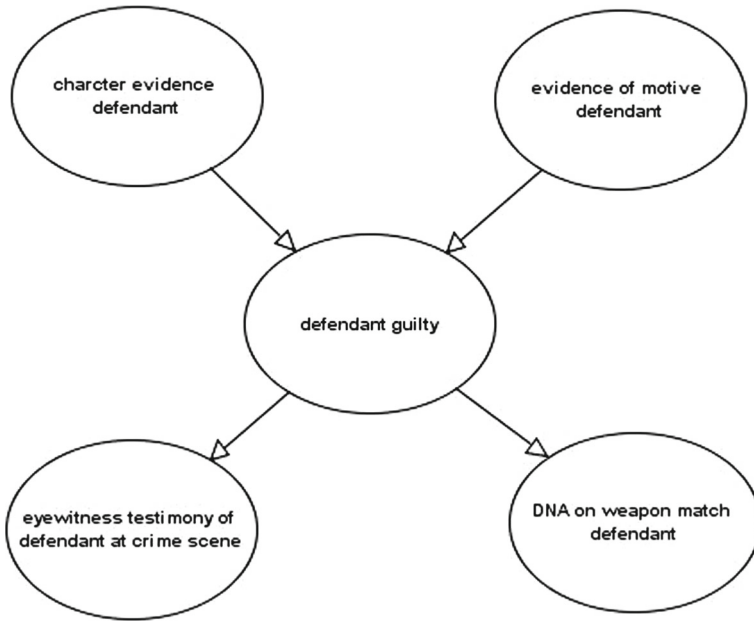


Fig. 6 Causal model of legal evidence (simplified structure without intermediate nodes)

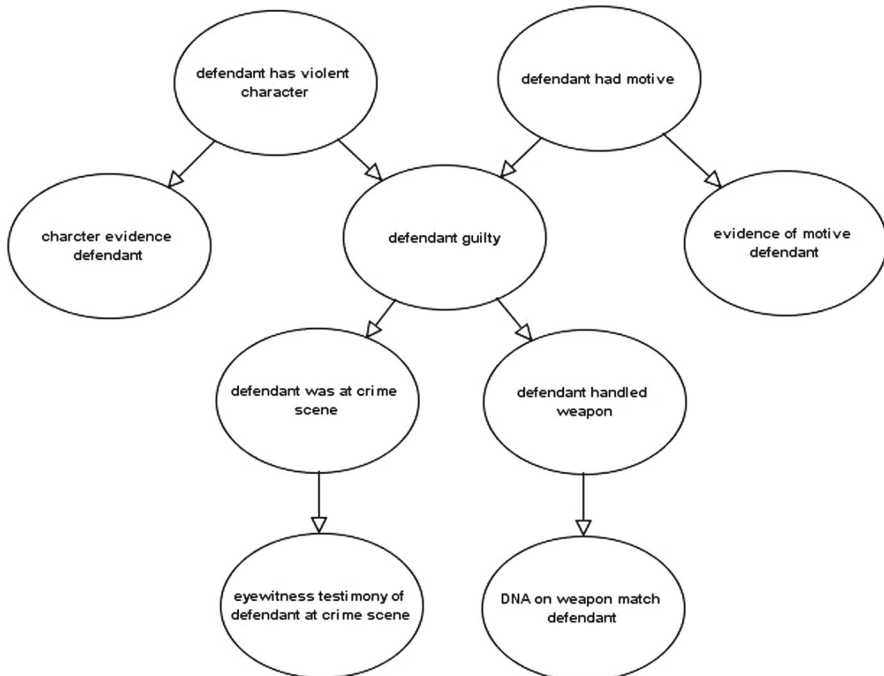


Fig. 7 Causal model of legal evidence (with some intermediate nodes)

the probability that the defendant is guilty given the character evidence depends on three factors: the probability that the defendant is guilty if he has a violent character, the prior probability that he has a violent character, and the likelihood of character evidence with regard to the hypothesis that the defendant has a violent character. The network helps us to calculate $P(\text{defendant guilty} | \text{character evidence defendant})$ given these factors.

The network can be elaborated further, but for now Fig. 7 provides a decent picture of a causal model. If, instead, the Bayesian networks is built as a reason model, the guilt hypothesis is modeled as a parent to all the evidence, since it is the *factum probandum* in relation to each and every one of them. See simplified structure without intermediate nodes (Fig. 8) and structure with some intermediate nodes (Fig. 9).

Reason models bear close resemblance to argument diagrams, for example so-called *Wigmore Charts* for legal evidence, developed by John Henry Wigmore more than hundred years ago (Wigmore, 1913), and the argument schemes for legal evidence

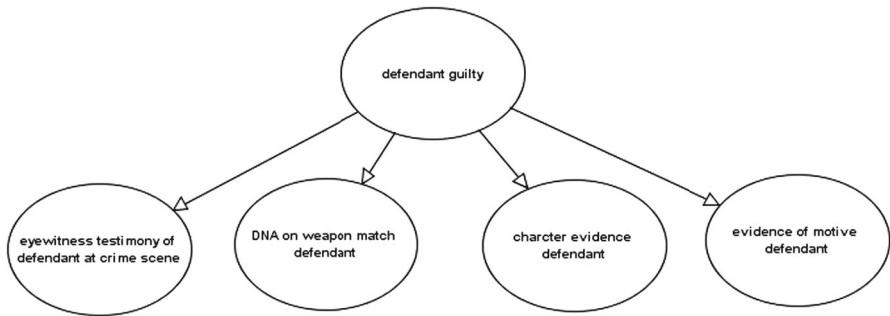


Fig. 8 Reason model of legal evidence (simplified structure without intermediate nodes)

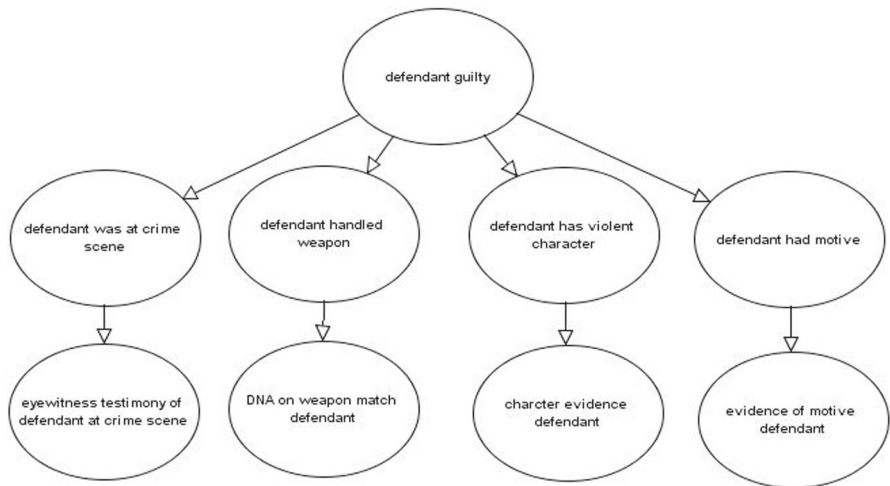


Fig. 9 Reason model of legal evidence (structure with some intermediate nodes)

developed in the last decades by Douglas Walton, Henry Prakken, Floris Bex and others (Bex, 2011, 2021; Bex et al., 2003; Walton, 2002). A striking difference, though, between reason models that use Bayesian networks and argument diagrams is that arrows in the latter point in the opposite direction, from *factum probans* to *factum probandum* to emphasize that reasoning starts with premises (or evidence) and moves towards the conclusion. When Bayesian networks are used to build reason models, the arrows point from *factum probandum* to *factum probans*, as we have seen, since Bayesian networks have the arrows point at the node where the probability table is conditioned on the other node.

Causal models, on the other hand, do not correspond to argument diagrams. With its emphasis on chronology and causal chains, causal modeling is closer to the scenario (or ‘story’) approach in the theory of legal evidence. Stories and arguments are quite different kinds of structures – the former are chains of events, while the latter consists of inferential steps. For this reason, it is just to be expected that the causal models and reason models will sometimes give rise to different networks, as in our example above.

In spite of these differences, there are many similarities between causal models and reason models, since both forms of modeling follow Bayesian probability theory. Let us for example consider a case where a glove belonging to X has been found at the crime scene, and this is evidence for the hypothesis that X is guilty but could also be explained by the alternative hypothesis that the glove was planted by the police (Fig. 10). In this case, the effect that the glove has on raising the probability of the hypothesis that X is guilty will be undercut if there is evidence raising the probability that the glove was planted by the police. This kind of propagation in a Bayesian network is known as ‘explaining away’ (Wellman & Henrion, 1993), and appears in the same way in causal modeling and reason modeling since it follows from Bayes’ Rule. Increasing the probability that the glove was planted by the police raises the probability of the false positive $P(\text{X glove at crime scene} = \text{True} | \text{X guilty} = \text{False})$, which lowers the likelihood ratio of the glove evidence with regard to the guilty-hypothesis. A Bayesian

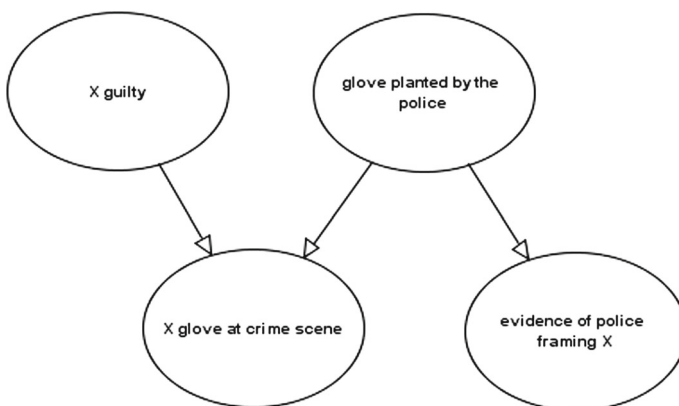


Fig. 10 Explaining away

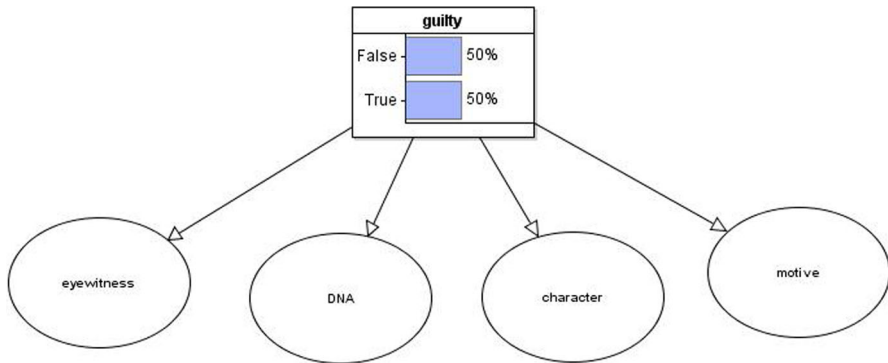


Fig. 11 Setting the prior at 50% to measure combined support

Network has the same structure (Fig. 10) whether it is constructed as a causal model or a reason model.

The purpose of reason models is to structure and assess the support that the evidence (*factum probans*) gives to the hypothesis at issue (*factum probandum*). Reason models are therefore better suited and more practical than causal models if we want to *measure the combined support* that the evidence gives to the hypothesis-at-issue in terms of likelihood. As Bayes' rule dictates, the prior odds of the hypothesis multiplied by the likelihood ratio equals the posterior odds of the hypothesis. The combined likelihood ratio for the evidence vis-à-vis the hypothesis can therefore be calculated by dividing the posterior odds in a reason model by the prior odds. To make this simple, the prior probability in the guilty-node can be set at 50%, making division by the prior odds a division by 1, and the combined likelihood ratio equal to the posterior probability that the hypothesis is true divided by the posterior probability that the hypothesis is false (Fig. 11).

Say, for example, that the prior probability in a reason model has been set at 50%, and the posterior probability calculated by the model, when all evidence nodes have been instantiated, is 95%. In this model, the guilt-node 'measures' the combined support of the evidence, since the odds of the guilt-node takes the same value as the combined likelihood ratio of the evidence: $0.95/0.05 = 19$ (Fig. 12).

This little trick cannot be used in causal models where the guilty-node is not a parent and its node probability table therefore does not have a prior that can be set to 50%.

It should be clarified that reason models do not preclude causal *thinking* in the construction of Bayesian networks for legal evidence. Causal thinking is helpful for identifying factors that should be included in a reason model. Before we build a Bayesian network, we must first identify the factors that will become nodes in the network. It is important that all relevant factors are found, and causal thinking is very helpful in this process. When we have evidence, we try to explain it, which is the same thing as answering the question: "Why do we have this evidence?". Most often, the answers will be possible causes (or "causal explanations"). Causes are a guide

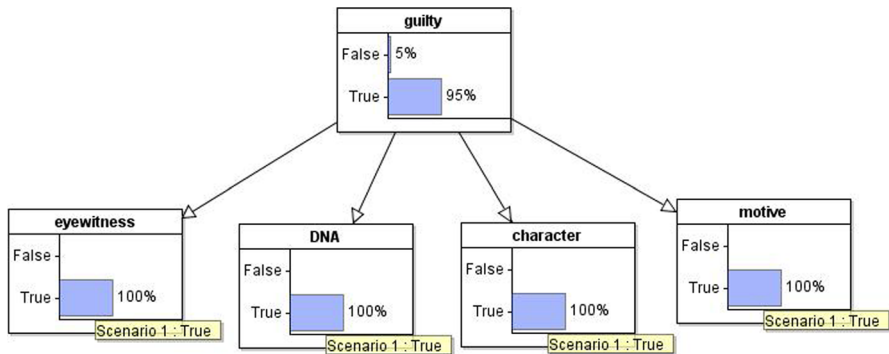


Fig. 12 Measuring combined support

to inferences (and truth) and play a crucial role when we are looking for possible inferences from the evidence.

The epistemological distinction between *context of discovery* and *context of justification* can be applied here. The identification of relevant factors takes place in the context of discovery, while the modeling of the network belongs to the context of justification. Causal thinking is very useful in the context of discovery, but that does not mean that relations of justification should be represented with a causal model. On the contrary, the proper way to model justification is a reason model.

It should also be pointed out that reason models do not exclude that causality is considered when values are inserted in the node probability tables. If, for example, A is known to always cause B, then $P(B|A) = 1$. Moreover, in the absence of reliable statistics the estimation of likelihood ratios will often have to rely on considerations about causal relations. Reason models is not about ignoring causal relations in probability assessments.

Whether one uses causal modeling or reason modeling it is of course important to take care and make sure that the model captures the right set of probabilistic relations and dependencies.

4 The problem of the prior

Bayesian modeling is a tool for assessing how the probability of a certain hypothesis is affected by new evidence, in which a *prior probability* is updated to a posterior probability. How this prior probability should be established is a pressing issue in Bayesian methodology, often referred to as the ‘problem of the prior’. In Bayesian modeling of criminal evidence, the hypothesis-at-issue is the so-called guilt-hypothesis, and the prior probability must be established according to the legal rule known as the *presumption of innocence*. The presumption of innocence is related to the prior probability since it prescribes that the defendant shall be viewed as innocent *before the evidence has been presented*. One implication of this principle is that the fact-finder

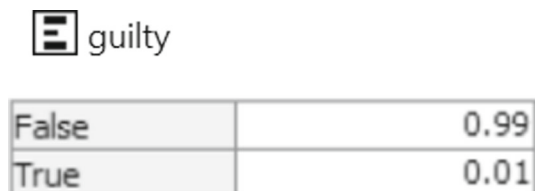
must ignore any assumption about the guilt rate of defendants. Thus, even if the fact-finder believes that, say, 80% of all defendants are guilty, he or she cannot start out with a prior probability of 0.8 for the hypothesis that the defendant is guilty. Even if this belief is based on ample experience of the legal system, the legal fact-finder is forbidden by the presumption of innocence to use this number as a prior probability of guilt. Such a prior would stereotype the defendant, and make it too easy to convict innocent defendants ().

Bayesian scholars have often suggested that the prior probability should be set at $1/N$, where N is the number of ‘possible perpetrators’. If, for example, a crime has been committed in a ‘closed room’ with 100 possible perpetrators including the defendant, the prior probability that the defendant did it, before any evidence pointing specifically at the defendant has been presented, should be taken to be $1/100 = 0.01$.

In a case where a crime has been committed in a ‘closed room’ the number of possible perpetrators is straightforward, but very few crimes in real life are committed under such circumstances. In cases where a crime is committed in the street, in the woods, or some other location that is not a closed room, it becomes debatable how the group of ‘possible perpetrators’ should be defined and delimited. Should it be limited to the inhabitants of the area where the crime took place? Or all inhabitants of the country where it took place? Or should it include the entire population of the world? How the presumption of innocence should be interpreted with regard to ‘possible perpetrators’ is a highly debated issue among legal scholars. For an overview of the literature see Dahlman (2017a) and Dahlman and Kolflaath (2021). In our discussion below we stick to the simple ‘closed room’ scenario and leave aside the problem of delimiting the number of possible perpetrators in other circumstances, as well as possible implications of the presumption of innocence for the prior probability of guilt. For a solution to the problem of the prior outside the ‘closed room’ scenario see Fenton et al. (2019).

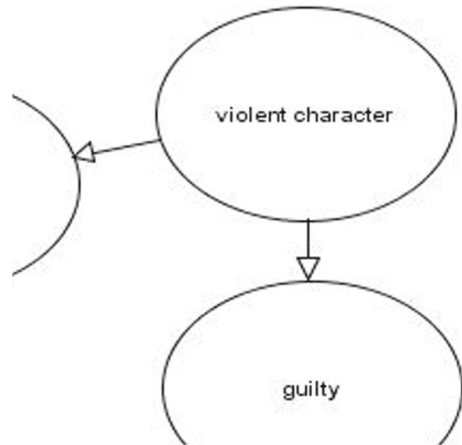
In a Bayesian network that is constructed as a reason model, a prior probability that reflects the number of possible perpetrators can easily be accommodated. In a reason model, the guilty-node is a parent to all other nodes connected to it, and the probability of guilt in its probability table will therefore always be a prior probability, unconditioned by other hypotheses in the case. The probability that the defendant is guilty before any other fact than the number of possible perpetrators has been established can therefore simply be inserted as a prior in the node probability table. In the closed-room case with 100 possible perpetrators, we can simply assign the value 0.01 to $P(\text{guilty} = \text{True})$ in the guilty node (Fig. 13).

Fig. 13 Prior probability in a reason model



False	0.99
True	0.01

Fig. 14 Guilty-node in a causal model



In a Bayesian network constructed as a causal model, where the guilty-node is a child to one or more nodes, a prior probability reflecting the number of possible perpetrators cannot be set in this way. Let us, for example, consider a causal model where the hypothesis that the defendant has a ‘violent character’ has been modeled as a parent to the guilt node (Fig. 14).

In this causal model, the probability that the defendant is guilty before other facts than the number of possible perpetrators has been established is not represented in the probability table of the guilty-node, since it no longer entails a ‘clean’ prior, but a prior conditioned on the hypothesis that the defendant has a violent character. And it cannot be represented in the violent-character-node either. The latter entails a clean, un-conditioned prior, yet it is not the prior probability that the defendant is guilty, but the prior probability that the defendant has a violent character.

Certainly, the number of possible perpetrators can be accommodated in a causal model, but it becomes a bit convoluted in comparison to the prior in the reason model. In our closed-room case the defendant is one of a hundred possible perpetrators. A prior probability of 1/100 can be accommodated in a causal model only if this probability is taken into account in combination with the prior for violent character in $P(\text{guilty} = \text{True} | \text{violent character} = \text{True})$. Let us, for example, assume that the prior probability that a random person has ‘violent character’ is 10%, and we therefore insert this value as a prior in the violent-character node. Let us also, for simplicity, assume that $P(\text{guilty} = \text{True} | \text{violent character} = \text{False}) = 0$. What value should be assigned to $P(\text{guilty} = \text{True} | \text{violent character} = \text{True})$? This probability should be set at 10%, if we consider that the prior probability of guilt is 1% ($0.1 = 0.01/0.1$), since $P(\text{guilty} | \text{violent character}) = P(\text{guilty}) / P(\text{violent character})$. If the defendant is one of the hundred persons in the closed room, and one in ten people have a violent character, there is a 10% probability that the defendant is one of the ten violent persons who could have committed the crime. Inserting these probabilities in the causal model (Fig. 15), is equivalent to the prior of 1% in the guilty-node of the proof model (Fig. 14).

To make a long story short, the number of possible perpetrators can be accommodated in causal modeling, but it becomes more complicated in comparison to a reason

Fig. 15 Prior probability in causal modeling

		violent character	
		False	True
False			0.9
True			0.1

		guilty	
		False	True
violent cha...	False		
False		1.0	0.9
True		0.0	0.1

model. This is clearly an advantage of reason models. There is an obvious risk that a legal fact-finder who uses a causal model will forget to incorporate the prior probability of guilt when assessing the probability of guilt given violent character, the result being that he or she ends up with a largely incorrect probability of guilt. $P(\text{guilt} = \text{True} | \text{violent character} = \text{True})$ can easily be confused with the predictive probability that a person with violent character will commit a certain violent act, that does not consider the number of possible perpetrators of the act at issue.

It should be mentioned, though, that in a reason model the number of possible perpetrators must instead be considered in the violent-character-node, modeled as a child to the guilty-node. The probability that the defendant has violent character given that he is innocent, $P(\text{violent character} = \text{True} | \text{guilty} = \text{False})$, should take the number of possible perpetrators into account. In our closed-room case, this probability is $9/99$. So, an advocate of causal models could say that there is no advantage with reason models with regard to the risk of error: whether you model in one way or the other, the number of possible perpetrators needs to be taken into account somewhere in the network, and there is always a risk that the fact-finder will forget to take the number of possible perpetrators into account and get the probability of guilt wrong. This is true, but such an error will typically have greater consequences in a causal model. An assessment of $P(\text{guilt} = \text{True} | \text{violent character} = \text{True})$ that does not consider the number of possible perpetrators can land wildly off the mark, but an assessment of $P(\text{violent character} = \text{True} | \text{guilty} = \text{False})$ that incorrectly equates this probability with $P(\text{violent character})$ will only err marginally, as long as the number of possible perpetrators is not extremely low. In our closed-room case, with 100 possible perpetrators, $P(\text{violent character} = \text{True} | \text{guilty} = \text{False}) = 9/99 \approx 0.09$ and $P(\text{violent character}) = 0.10$. To our knowledge it has not been investigated empirically to what extent fact-finders make these errors, so the practical consequences of the difference between causal modeling and reason modeling with regard to the problem of the prior remains as a question for future research.***

5 Conclusion

In this paper we have compared causal models with reason models in the construction of Bayesian network for legal evidence. In causal models arrows in the network are drawn from causes to effects. In a reason models the arrows are instead drawn towards the evidence, from *factum probandum* to *factum probans*. We have explored the differences between causal models and reason models and observed several distinct advantages with reason models.

- Reason models are better aligned with the philosophy of Bayesian inference, as they model reasons for up-dating beliefs (Sect. 1).
- Reason models are better suited for measuring the combined support of the evidence vis-à-vis the guilt-hypothesis in terms of likelihood (Sect. 3).
- A prior probability of guilt that reflects the number of possible perpetrators is accommodated more easily with reason models (Sect. 4).

Maybe advocates of causal models will come up with counterexamples and argue that causal models have some distinct advantages over reason models in the context of legal evidence. We are looking forward to this conversation.

Funding Open access funding provided by Lund University. Funding was provided by Torsten Söderbergs Stiftelse and Ragnar Söderbergs stiftelse.

Declarations

Conflict of interest All authors declare that there are no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bex, F. (2011). *Arguments, stories and criminal evidence—A formal hybrid theory*. Springer.
- Bex, F. (2021). Argumentation and evidence. In C. Dahlman, A. Stein, & G. Tuzet (Eds.), *Philosophical foundations of evidence law*. Oxford University Press.
- Bex, F., Prakken, H., Reed, C., & Walton, D. (2003). Towards a formal account of reasoning about evidence—argumentation schemes and generalizations. *Artificial Intelligence and Law*, 11, 125–165.
- Dahlman, C. (2017a). Determining the base rate for guilt. *Law, Probability and Risk*, 17, 15–28.
- Dahlman, C. (2017b). Unacceptable generalizations in arguments on legal evidence. *Argumentation*, 31, 83–99.
- Dahlman, C., & Kolfaath, E. (2021). The problem of the prior in criminal trials. In C. Dahlman, A. Stein, & G. Tuzet (Eds.), *Philosophical foundations of evidence law*. Oxford University Press.

- Fenton, N., & Neil, M. (2018). *Risk assessment and decision analysis with Bayesian networks* (2nd ed.). CRC Press.
- Fenton, N., Neil, M., & Lagnado, D. (2013). A general structure for legal arguments about evidence using Bayesian networks. *Cognitive Science*, 37, 61–102.
- Fenton, N., Lagnado, D., Dahlman, C., & Neil, M. (2019). The opportunity prior—A proof based prior for criminal cases. *Law, Probability and Risk*, 18, 237–253.
- Lagnado, D. (2021). *Explaining the evidence*. Cambridge University Press.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. Morgan Kaufmann.
- Pearl, J. (2009). *Causality—Models, reasoning and inference* (2nd ed.). Cambridge University Press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *causal inference in statistics—A primer*. Wiley.
- Schweizer, M. (2015). *Beweiswürdigung und Beweismaß*. Mohr Siebeck.
- Taroni, F., Aitken, C., Garbolino, P., & Biedermann, A. (2006). *Bayesian networks and probabilistic inference in forensic science*. Wiley.
- Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law*, 24, 285–324.
- Walton, D. (2002). *Legal argumentation and evidence*. Penn State University Press.
- Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge University Press.
- Wellman, M., & Henrion, M. (1993). Explaining ‘explaining away.’ *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, 287–292.
- Wigmore, J. H. (1913). The problem of proof. *Illinois Law Review*, 8, 77–103.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.