Awareness Growth with Bayesian Networks

March 20, 2023

Wordcount (including footnotes): 5,392

We examine different counterexamples to Reverse Bayesianism, a popular theory that addresses the problem of awareness growth. We agree with the general skepticism toward Reverse Bayesianism, but submit that the problem of awareness growth cannot be tackled in an algorithmic manner, because subject-matter, structural assumptions need to be made explicit and play a role in determining how probabilities should be updated in light of awareness growth. Thanks to their ability to express probabilistic dependencies, we illustrate how Bayesian networks can help to model awareness growth in the Bayesian framework.

1 Introduction

Learning is modeled in the Bayesian framework by the rule of conditionalization. This rule posits that the agent's new degree of belief in a proposition H after a learning experience E should be the same as the agent's old degree of belief in H conditional on E. That is,

$$\mathsf{P}^E(H) = \mathsf{P}(H|E),$$

where $\mathsf{P}()$ represents the agent's old degree of belief (before the learning experience E) and $\mathsf{P}^E()$ represents the agent's new degree of belief (after the learning experience E).

Both E and H belong to the agent's algebra of propositions. This algebra models the agent's awareness state, the propositions taken to be live possibilities. Conditionalization never modifies the algebra and thus makes it impossible for an agent to learn something they have never thought about. Even before learning about E, the agent must already have assigned a degree of belief to any proposition conditional on E. This picture commits the agent to the specification of their 'total possible future experience' (Howson, 1976), as though learning was confined to an 'initial prison' (Lakatos, 1968).

But, arguably, the learning process is more complex than what conditionalization allows. Not only do we learn that some propositions that we were entertaining are true or false, but we may also learn new propositions that we did not entertain before. Or we may entertain new propositions—without necessarily learning that they are true or false—and this change in awareness may in turn change what we already believe. How should this more complex learning process be modeled by Bayesianism? Call this the problem of awareness growth.

The algebra of propositions need not be so narrowly construed that it only contains propositions that are presently under consideration. The algebra may also contain propositions which, though outside the agent's present consideration, are still the object, perhaps implicitly, of certain dispositions to believe.[1] But even this expanded algebra will have to be revised sooner or later. The algebra of propositions could in principle contain anything that could possibly be conceived, expressed, thought of. Such a rich algebra would not need to change at any point, but this is an implausible model of ordinary agents with bounded resources such as ourselves.

Critics of Bayesianism and sympathizers alike have been discussing the problem of awareness growth under different names for quite some time, at least since the eighties. This problem arises in a number of different contexts, for example, new scientific theories (Chihara, 1987; Earman, 1992; Glymour, 1980), language changes and paradigm shifts (Williamson, 2003), and theories of induction (Zabell, 1992). A proposal that has attracted considerable scholarly attention in recent years is Reverse Bayesianism (Bradley, 2017; Karni & Vierø, 2015; Wenmackers & Romeijn, 2016). The idea is to model awareness growth as a change in the algebra while ensuring that the proportions of probabilities of the propositions shared between the old and new algebra remain the same in a sense to be specified.

Let $\mathscr{F}$ be the initial algebra of propositions and let $\mathscr{F}^+$ be the algebra after the agent's awareness state has grown. Both algebras contain the contradictory and tautologous propositions $\bot$ and $\top$, and they are closed under connectives such as disjunction $\vee$, conjunction $\wedge$ and negation $\neg$. Denote by X and X^+ the subsets of these algebras that contain only basic propositions, namely those without connectives. **Reverse Bayesianism** posits that the ratio

[1] Roussos (2021) notes that, for the sake of clarity, the problem of awareness growth should only address propositions which agents are *truly* unaware of (say new scientific theories), not propositions that were temporarily forgotten or set aside. This is a helpful clarification to keep in mind, although the recent literature on the topic does not make a sharp a distinction between true unawareness and temporary unawareness.

of probabilities for any basic propositions A and B in both X and X^+—the basic propositions shared by the old and new algebra—remain constant through the process of awareness growth:

$$\frac{\mathsf{P}(A)}{\mathsf{P}(B)} = \frac{\mathsf{P}^+(A)}{\mathsf{P}^+(B)},$$

where $\mathsf{P}()$ represents the agent's degree of belief before awareness growth and $\mathsf{P}^+()$ represents the agent's degree of belief after awareness growth.

Reverse Bayesianism is an elegant theory that manages to cope with a seemingly intractable problem. As the awareness state of an an agent grows, the agent would prefer not to throw away completely the epistemic work they have done previously. The agent may desire to retain as much of their old degrees of beliefs as possible. Reverse Bayesianism provides a simple recipe to do that. It also coheres with the conservative spirit of Bayesian conditionalization which preserves the old probability distribution conditional on what is learned.

Unfortunately, Reverse Bayesianism does not deliver the intuitive results in all cases. There is no shortage of counterexamples against it in the recent philosophical literature (Mathani, 2020; Steele & Stefánsson, 2021). In addition, attempts to extent traditional arguments in defense of Bayesian conditionalization to Reverse Bayesianism seem to hold little promise (Pettigrew, forthcoming). If the consensus in the literature is that Reverse Bayesianism is not the right theory of awareness growth, what theory (if any) should replace it?

Here we offer a diagnosis of what is wrong with Reverse Bayesianism and outline an alternative proposal. The problem of awareness growth—we hold—cannot be tackled in an algorithmic manner because subject-matter assumptions, both probabilistic and structural, need to be made explicit. So any theory of awareness growth cannot be a purely formal theory. This does not mean, however, we should give up on probability theory altogether. Thanks to their ability to express probabilistic dependencies, Bayesian networks can help to model awareness growth in the Bayesian framework. We illustrate this claim as we examine different counterexamples to Reverse Bayesianism.

The plan for the paper is as follows. To set the stage for the discussion, Section 2 begins with two counter-exmaples to Reverse Bayesianism by Steele & Stefánsson (2021). One example targets awareness expansion and the other awareness refinement (more on the distinction soon). Section 3 models cases of awareness expansion using Bayesian networks. Section 4 further illustrates the fruitfulness of this approach by looking at two counter-examples to Reverse Bayesianism by Mathani (2020). Section 5 turns to awareness refinement. Finally, Section 6 outlines a general theory of awareness growth with Bayesian networks.

2 Steele and Stefánsson's Examples

In this section, we rehearse two of the counterexamples to Reverse Bayesianism by Steele & Stefánsson (2021). One example targets awareness expansion and the other awareness refinement. A precise definition of expansion can be tricky to provide, but a rough characterization will suffice for now. Suppose, as is customary, propositions are interpreted as sets of possible worlds, where the set of all possible worlds is the possibility space. Awareness expansion occurs when a new proposition is added to the algebra and its interpretation includes possible worlds not in the original possibility space. So the addition of the new proposition causes the possibility space to expand. By contrast, awareness refinement (roughly) occurs when the new proposition added to the algebra induces a more fine-grained partition of the possibility space.

The most straightforward case of *awareness expansion* occurs when you become aware of a new explanation for the evidence at your disposal which you had not considered before. This can happen in many fields of inquiry: medicine, law, science, everyday affairs. Here is a scenario by Steele & Stefánsson (2021):

> FRIENDS: "Suppose you happen to see your partner enter your best friend's house on an evening when your partner had told you she would have to work late. At that point, you become convinced that your partner and best friend are having an affair, as opposed to their being warm friends or mere acquaintances. You discuss your suspicion with another friend of yours, who points out that perhaps they were meeting to plan a surprise party to celebrate your upcoming birthday—a possibility that you had not even entertained." [Sec. 5, Example 2]

Initially, the algebra contained the hypotheses 'my partner and my best friend met to have an affair' (*affair*) and 'my partner and my best friend met as friends' (*friends*). These were the only explanations you considered for the fact that your partner and your best friend met one night without telling you. But, when the algebra changes, a new hypothesis is added which you had not considered before: your partner and your best friends met to plan a surprise party for your upcoming birthday (*surprise*).

This change in the algebra is not innocuous. At first, the hypothesis *affair* seems more likely than *friends* because the former seems a better explanation than the latter for the secretive behavior of your partner and best friend.[2] But when the new hypothesis *surprise* is added, things change: *surprise* now seems a better explanation overall and thus more likely than *affair*. And since *surprise* implies *friends*, the latter must be more likely than *affair*. This conclusion violates Reverse Bayesianism since the ratio of the probabilities of *friends* and *affair* has changed before and after awareness expansion:

$$\frac{\mathsf{P}(\mathit{affair})}{\mathsf{P}(\mathit{friends})} > 1 \text{ but } \frac{\mathsf{P}^{+}(\mathit{affair})}{\mathsf{P}^{+}(\mathit{friends})} < 1.$$

Steele & Stefánsson note that a quick fix is available. It is reasonable to suppose that no change in the probabilities should occur so long as we confine ourselves to the old probability space. With this in mind, consider the following condition, called **Awareness Rigidity**:

$$\mathsf{P}^{+}(A|T^{*}) = \mathsf{P}(A),$$

where T^{*} corresponds to a proposition that picks out, from the vantage point of the new awareness state, the entire possibility space *before* the episode of awareness growth. Awareness rigidity posits that, once a suitable proposition T^{*} is identified, the old probability assignments remain unchanged conditional on T^{*}. In our running example, $\neg$*surprise* is the suitable proposition T^{*}: *that* there was no surprise party in the making picks out the original possibility space. Conditional on $\neg$*surprise*, no probability assignment should change, including the probability of *affair*. This is the intended result.

But this is not the end of the story. Steele & Stefánsson go on to show that Awareness Rigidity does not hold in other cases, what they call *awareness refinement*. As noted before, these are cases in which the new proposition induces a more fine-grained partition of the possibility space. Consider this scenario:

> MOVIES: "Suppose you are deciding whether to see a movie at your local cinema. You know that the movie's predominant language and genre will affect your

[2] This assumes that the prior probabilities of the two hypotheses were not strongly skewed in one direction. If you were initially nearly certain your partner could not possibly have an affair with your best friend, the fact they behaved secretively or lied to you should not affect that much the probability of the two hypotheses.

viewing experience. The possible languages you consider are French and German and the genres you consider are thriller and comedy. But then you realise that, due to your poor French and German skills, your enjoyment of the movie will also depend on the level of difficulty of the language. Since it occurs to you that the owner of the cinema is quite simple-minded, you are, after this realisation, much more confident that the movie will have low-level language than high-level language. Moreover, since you associate low-level language with thrillers, this makes you more confident than you were before that the movie on offer is a thriller as opposed to a comedy." (Steele & Stefánsson, 2021, sec. 5, Example 3)

You initially categorized movies by just language and genre, and then you refined your categorization by adding another variable, level of difficulty. Without considering language difficulty, you assigned the same probability to the hypotheses *thriller* and *comedy*. But learning that the owner was simple-minded made you think that the level of linguistic difficulty must be low and the movie most likely a thriller rather than a comedy (perhaps because thrillers are simpler—linguistically—than comedies). Since the probability of *thriller* goes up, this scenario violates (against Reverse Bayesianism) the condition $\frac{\mathsf{P}(thriller)}{\mathsf{P}(comedy)} = \frac{\mathsf{P}^+(thriller)}{\mathsf{P}^+(Comedy)}$. For the same reason, it also violates (against Awareness Rigidity) the condition $\mathsf{P}(thriller) = \mathsf{P}^+(thriller|thriller \vee comedy)$, where $thriller \vee comedy$ is a proposition that picks out the entire possibility space.[3]

Some might object that the probability of *thriller* goes up, not because of awareness refinement, but because you learn that the owner is simple-minded. And if learning in the strict Bayesian sense—one modeled by conditionalization—takes place, it should be no surprise that some probabilities will change. Admittedly, the scenario can be split into two sub-episodes. In the first, you entertain a new variable besides language and genre, namely the language difficulty of the movie. In the second episode, you learn something new, namely that the owner is simple-minded. Arguably, the violation of Reverse Bayesianism or Awareness Rigidity can be attributed to learning that the owner is simple-minded, not to the mere growth in awareness. But there are cases of awareness refinement that do no involve learning in the Bayesian sense and still violate Reverse Bayesianism and Awareness Rigidity (see later in Section 5). So it is incumbent to understand under what circumstances these principles fail.

As will become clear, we believe that theorizing about awareness growth should be grounded

[3] Since MOVIES is a case of refinement, $thriller \vee comedy$ picks out the entire possibility space both before and after awareness growth.

in the subject-matter information underlying the scenario at hand. This subject-matter takes many forms. In FRIENDS, awareness expansion does not change the basic presupposition that someone's behavior must have a reason. In MOVIES, awareness refinement does not change the fact that characteristics such as language, difficulty or genre may influence one's decision to select a movie for showing rather than another. What is wrong with principles such as Reverse Bayesianism or Awareness Rigidity—we hold—is that they are purely formal. We need instead a formalism that can represent the relevant subject-matter information. In what follows, we illustrate how Bayesian networks can serve this purpose. We first consider expansion (Section 3 and 4) and then refinement (Section 5).

3 Expansion with Bayesian Networks

A Bayesian network is a compact formalism to represent probabilistic dependencies. It consists of a direct acyclic graph (DAG) accompanied by a probability distribution. The nodes in the graph represent random variables that can take different values. We will use 'nodes' and 'variables' interchangeably. The nodes are connected by arrows, but no loops are allowed, hence the name 'direct acyclic graph'. A simple graph structure we will use repeatedly is the so-called *hypothesis-evidence idiom* (Fenton, Neil, & Lagnado, 2013):

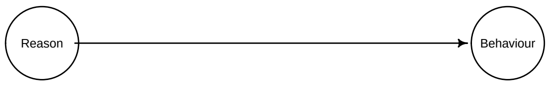
$$H \longrightarrow E$$

where H is the hypothesis node (upstream) and E the evidence node (downstream). If an arrow goes from H to E, the full probability distribution associated with the Bayesian network is defined by two probability tables.[4] One table defines the prior probabilities $\mathsf{P}(H = h)$ for all the states (or values) of H, and another table defines the conditional probabilities of the form $\mathsf{P}(E = e|H = h)$, where uppercase letters represent the variables (nodes) and lower case letters represent the states (or values) of these variables. These two probability tables are sufficient to specify the full probability distribution. The other probabilities—say $\mathsf{P}(E = e)$, $\mathsf{P}(H = h|E = e)$, etc.—follow by simply applying the probability axioms. As nedded, more complex graphical structures can also be used.

[4]A major point of contention in the interpretation of Bayesian networks is is the meaning of the directed arrows. They could be interpreted causally—as though the direction of causality proceeds from the events described by the hypothesis to event described by the evidence—but they need not be; see footnote 18.

Bayesian networks are relied upon in many fields, but have been rarely deployed to model awareness growth (the exception is Williamson (2003)). We think instead they are a good framework for this purpose. Awareness growth can be modeled as a change in the graphical network—nodes and arrows are changed, added or erased—as well as a change in the probability distribution from the old to the new network. In this section, we focus on cases of awareness expansion and turn to refinement in the next section.

Recall the scenario FRIENDS from before. It can be modeled with the hypothesis-evidence idiom. The graph can be made more perspicuous by labeling the downstream node 'Behavior' (the evidence or fact to be explained) and the upstream node 'Reason' (the explanation or hypothesis about the cause of the behavior):



Initially, before awareness growth, the hypothesis node *Reason* takes only two states: *friends* and *affair*. These two states are meant to be exclusive and exhaustive, so *affair* functions as the negation of *friends*, and vice versa. After awareness growth—specifically, awareness expansion—the two states are no longer exhaustive. A third state is added: *surprise*. So, in the formalism we are using, expansion simply consists in the addition of an extra state to one of the nodes of the network. The rest of the structure of the network remains intact.

Recall that the ratio of posterior probabilities of *friends* to *affairs* changed as a result of awareness expansion. The fact that your partner and best fried met without telling you—call this behavior *secretive*—initially made *affair* more likely than *friends*, but then the same fact made *friends* more likely than *affair* after awareness growth. So, formally,

$$\frac{\mathsf{P}(Reason{=}affair|Behavior{=}secretive)}{\mathsf{P}(Reason{=}friends|Behavior{=}secretive)} > 1 \text{ but } \frac{\mathsf{P}^{+}(Reason{=}friends|Behavior = secretive)}{\mathsf{P}^{+}(Reason{=}affair|Behavior{=}secretive)} > 1.$$

Despite this change in posterior probabilities, it is plausible to assume that the likelihoods do not change. Even after awareness expansion, the fact that your partner and best fried met without telling you—*secretive*—makes better sense in light of *affair* compared to *friends*:

$$\frac{\mathsf{P}(Behavior{=}secretive|Reason{=}affair)}{\mathsf{P}(Behavior{=}secretive|Reason{=}friends)} = \frac{\mathsf{P}^{+}(Behavior{=}secretive|Reason{=}affair)}{\mathsf{P}^{+}(Behavior{=}secretive|Reason{=}friends)} > 1.$$

This equality holds even though the novel explanation *surprise* introduced during awareness expansion makes better sense of the secretive behavior overall:[5]

$$\frac{\mathsf{P}^{+}(\mathit{Behavior{=}secretive}|\mathit{Reason{=}surprise})}{\mathsf{P}^{+}(\mathit{Behavior{=}secretive}|\mathit{Reason{=}affair})} > 1.$$

This analysis suggests a slight reformulation of conditions such as Reverse Bayesianism and Awareness Rigidity. Instead of comparing posterior probabilities of hypotheses given the evidence ($\mathsf{P}(H = h|E = e)$), we can compare which explanation or hypothesis makes better sense of the evidence ($\mathsf{P}(E = e|H = h)$). So, for all values h, h' and e of upstream node H and downstream node E in the old network, consider the following constraint:

$$\frac{\mathsf{P}(E = e|H = h)}{\mathsf{P}(E = e|H = h')} = \frac{\mathsf{P}^{+}(E = e|H = h \,\&\, X \neq x^{*})}{\mathsf{P}^{+}(E = e|H = h' \,\&\, X \neq x^{*})}, \tag{C}$$

where x^{*} is the new state added and X is the node (upstream or downstream) to which the new state belongs, such as $H = \mathit{surprise}$ in FRIENDS.

In words, the constraint says that one's old assessment of the relative plausibility of two competing hypotheses in light of a fixed body of evidence should remain unchanged after awareness expansion. Constraint (C) is a variant of Reverse Bayesianism that only applies to conditional probabilities of the form $\mathsf{P}(E = e|H = h)$ for Bayesian networks of the form $H \to E$. In addition, the constraint mimics Awareness Rigidity in that it ensures that the conditional probabilities in question exclude the novel state $X = x^{*}$.[6]

How generally does constraint (C) apply besides examples such as FRIENDS? We put forward the following *working hypothesis*:

> If there is no change in the structure of the network, constraint (C) holds generally.
> If there is a change in the structure of the network, constraint (C) will fail under certain conditions (to be specified later).

Since expansion—as we have defined it—does not change the network structure, the constraint should always hold for expansion.[7] We provide support for this claim in the next section.

---

[5] Even though $\frac{\mathsf{P}^{+}(\mathit{Behavior{=}secretive}|\mathit{Reason{=}surprise})}{\mathsf{P}^{+}(\mathit{Behavior{=}secretive}|\mathit{Reason{=}affair})} > 1$ and $\frac{\mathsf{P}^{+}(\mathit{Behavior{=}secretive}|\mathit{Reason{=}affair})}{\mathsf{P}^{+}(\mathit{Behavior{=}secretive}|\mathit{Reason{=}friends})} > 1$—so *affair* still makes better sense of the evidence than *friends* before and after awareness expansion—the posterior probability of *friends* is higher than *affair* after awareness expansion.

[6] When the novel state is added to the upstream node H, the condition $X \neq x^{*}$ is redundant. We will see later cases in which the novel state is added to the downstream node E, and here the condition is not redundant.

[7] It is crucial that the new hypothesis or explanation does not change the existing structure of the network. Consider this example. You are wondering which horse will win the race. You have done a careful study of past performances under different conditions and concluded that Red is more likely to win than Green. But you

4 Mathani's Examples

To acquire a firmer grasp on constraint (C), we now examine a couple of examples by Mathani (2020). They are intended as counterexamples to Reverse Bayesianism, as well as a challenge to the distinction between expansion and refinement. When modeled with Bayesian networks, however, Mathani's examples are straightforward cases of awareness expansion.

The first example goes like this:

> TENANT: You are staying at Bob's flat which he shares with his landlord. You know that Bob is a tenant, and that there is only one landlord, and that this landlord also lives in the flat. In the morning you hear singing coming from the shower room, and you try to work out from the sounds who the singer could be. At this point you have two relevant propositions that you consider possible ... *landlord* standing for the possibility that the landlord is the singer, and *bob* standing for the possibility that Bob is the singer . . . Because you know that Bob is a tenant in the flat, you also have a credence in the proposition *tenant* that the singer is a tenant. Your credence in *tenant* is the same as your credence in *bob*, for given your state of awareness these two propositions are equivalent ... Now let's suppose the possibility suddenly occurs to you that there might be another tenant living in the same flat (*other*).

Initially, you thought the singer could either be the landlord or Bob, the tenant. Then you come to the realization that a third person could be the singer, another tenant. The possibility that there could be a third person in the shower—besides Bob or the landlord—is a novel explanation for why you hear singing in the shower. So TENANT seems to be a standard case of expansion like FRIENDS. At the same time, this scenario is a bit more complicated. The expansion in awareness goes along with an interesting conceptual shift. Before awareness expansion, that Bob is in the shower and that a tenant is in the shower are equivalent descriptions, but after the expansion, this equivalence breaks down.

have not considered the possibility that Grey would run. If Grey does run, it will have a greater chance of winning than the others, but will also make—for some odd reason—Green a much better racer than Red. So the odds that Green will win compared to Red should now be higher. Here, the new hypothesis introduces novel information that was not known before, say that the participation of Grey would weaken Red's performance and strengthen Green's performance. So the network should be changed in two ways: first, a new state should be added to the outcome node (Green wins; Red wins; Grey wins); and second, a new node should be added modeling the fact that Grey is participating and its participation affects Red's and Green's performance.

As Mathani shows, this scenario challenges Reverse Bayesianism. For it is natural to assign $1/3$ to $landlord$, bob and $other$ after awareness growth, and 1/2 to $landlord$ and bob before awareness growth. That someone is singing in the shower is evidence that someone must be in there, but without any more discriminating evidence, each person should be assigned the same probability. Consequently, a probability of 2/3 should be assigned to $tenant$ after awareness growth, but only 1/2 before. On this picture, the proportion of $landlord$ to $tenant$ changes from 1:1 (before awareness growth) to 1:2 (after awareness growth).[8, 9]

One could resist the challenge. For recall that Reverse Bayesianism only applies to basic propositions, which we defined earlier as propositions without connectives. So a possible fix is to adopt the following principle: if two propositions happen to be equivalent relative to some awareness state, they cannot be both considered basic. In TENANT, since *bob* and $tenant$ are initially equivalent descriptions of the same state of affairs, they would not be considered both basic propositions. If only *bob* is considered basic, along with $landlord$, then the proportion of the probability of *bob* and $landlord$ would remain the same during awareness growth, but not the proportion of the probabilities of $tenant$ and $landlord$. This yields the intuitive result. However, if $tenant$ is considered basic along with $landlord$, Reverse Bayesianism would require that the proportion of the probability of $tenant$ and $landlord$ remain the same during awareness growth. This is counterintuitive.

So dividing propositions into basic and non-basic is riddled with difficulties. There is no obvious way to draw the line between the two, except as a mere syntactic distinction. Still, the above discussion alerts us to the fact that a difference exists between propositions like *bob* and those like $tenant$. The latter describes a role that different people could play besides Bob. Bayesian networks can help to model the person/role distinction, as follows:

[8]Here is a more involved argument. Suppose, after you hear singing in the shower, you become sure someone is in there, but you cannot tell who. So $\mathsf{P}(landlord) = \mathsf{P}(bob) = 1/2$, and since *bob* and *tenant* are equivalent, also $\mathsf{P}(tenant)$ = 1/2. Now, *landlord*, *Bob* and *tenant* are all propositions that you were originally aware of, and thus Reverse Bayesianisn requires that their probabilities should remain in the same proportion after your awareness grows. But note that *other* entails $tenant$ and *bob* and *Other* are disjoint, so it follows that $\mathsf{P}^+(other)$ must have zero probability. If $\mathsf{P}^+(other)$ were greater than zero, the proportion of of the probability of $tenant$ to *landlord* (or the proportion of the probability of *bob* to *landlord*) should change.

[9]This scenario need not be a challenge for Awareness Rigidity. Much depends on the choice of the proposition T^* that picks out, from the vantage point of the new awareness state, the old possibility space prior to awareness growth. The proposition $landlord \vee bob$ does the job. For $\mathsf{P}^+(landlord|landlord \vee bob)$ and $\mathsf{P}^+(bob|landlord \vee bob)$ should both equal $1/2$, and thus $\mathsf{P}^+(other|landlord \vee bob) = 0$, but this does not mean that $\mathsf{P}^+(other|landlord \vee tenant)$ should equal zero. This is the intended result.

$\mathsf{P}(Role \mid Person)$		*Person*	
		landlord-person	*bob*
Role	*tenant*	0	1
	landlord	1	0
	TOTAL	1	1

$\mathsf{P}^{+}(Role \mid Person)$		*Person*		
		landlord-person	*bob*	*other*
Role	*tenant*	0	1	1
	landlord	1	0	0
	TOTAL	1	1	1

$\mathsf{P}(Person)$	*Person*	
	landlord-person	*bob*
	1/2	1/2

$\mathsf{P}^{+}(Person)$	*Person*		
	landlord-person	*bob*	*other*
	1/3	1/3	1/3

Table 1: The table displays a plausible probability distribution for the TENANT scenarios. Constraint (C) is met.

This subject-matter information—the distinction between people and the role they play—remain fixed throughout the process of awareness expansion. What changes is how the details are filled in. Initially, the upstream node *Person* has two possible states, representing who could be in the bathroom singing: *landlord-person* and *bob*.[10] The downstream node *Role* has also two values, *landlord* and *tenant*. After your awareness grows, the upstream node *Person* should now have one more possible state, *other*. Crucially, note that since TENANT is a case of expansion by our definition—a state was added to a node—constraint (C) should hold. This is precisely what happens. After all, the conditional probabilities $\mathsf{P}(Role = landlord \mid Person = landlord\text{-}person)$ and $\mathsf{P}(Role = landlord \mid Person = bob)$ do not change after awareness growth (see Table 1).

It is now time to test the tenability of constraint (C) in cases of awareness expansion a bit further. So far we only considered cases of expansion in which a new state was added to an *upstream* node. In FRIENDS, a state was added to the upstream node *Reason*, and in TENANT, a state was added to the upstream node *Person*. What if the new state was added to a downstream node? For consider a variation of FRIENDS. Suppose the downstream node *Behavior* could initially take only two values, say *secretive* and *transparent*. That is, the behavior of your

[10]To simplify things, the assumption here is that the evidence of singing has already ruled out the possibility that no one would be in the shower. In principle, the network should be more complex and contain another node for the evidence to be explained (the fact of singing in the shower), as follows: *Singing* $\leftarrow$ *Person* $\rightarrow$ *Role*.

partner and best friend could be secretive or transparent. You then realize there could be ambiguity in the behaviour, say the behavior could have elements of secrecy and elements of transparency. So the node could take a third value: *ambigous*. Initially, *secretive* and *transparent* are considered exhaustive, but that is no longer true after adding *ambiguous*. So the old conditional probabilities will change, specifically, $\mathsf{P}(E = e|H = h) \neq \mathsf{P}^+(E = e|H = h)$, where H is the upstream node and E downstream. However, if we exclude the novel state, there should be no change, so $\mathsf{P}(E = e|H = h \,\&\, E \neq e^*) = \mathsf{P}^+(E = e|H = h \,\&\, E \neq e^*)$, where e^* is the novel state added to the downstream node E. Hence, constraint (C) should be satisfied.

The same analysis applies to a more complex example by Mathani:

> COIN: You know that I am holding a fair ten pence UK coin which I am about to toss. You have a credence of 0.5 that it will land *heads*, and a credence of 0.5 that it will land *tails*. You think that the tails side always shows an engraving of a lion. So you also have a credence of 0.5 that it will land with the lion engraving face-up (*lion*): relative to your state of awareness *tails* and *lion* are equivalent.... Now let's suppose that you somehow become aware that occasionally ten pence coins have .... an engraving of Stonehenge on the tails side (*stonehenge*).

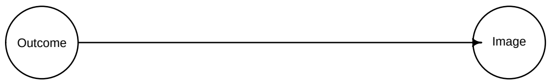
The propositions *tails* and *lion* are equivalent prior to awareness growth. Suppose you initially gave *tails* and *lion* the same credence. If they are basic propositions, Reverse Bayesianism would require that their relative probabilities should stay the same after awareness grow. The same is true of *heads* and *tails*. But since *lion* and *stonehenge* are incompatible and the latter entails *tails*, you should have $\mathsf{P}^+(stonehenge) = 0$, an undesirable conclusion.

Mathani observes that this scenario blurs the distinction between expansion and refinement. For one thing, COIN seems a case of refinement. The space of possibilities is held fixed—the coin could come up heads or tails—but the options for tails are further refined, for tails could be *lion* or *stonehenge*. On the other hand, a new possibility has been added after awareness growth, namely *stonehenge*, which had not been considered before. This would indicate that COIN is a case of expansion. This ambiguity makes it difficult to settle whether the scenario is a challenge for Awareness Rigidity.[11]

[11] If it is a case of refinement, $heads \vee tails$ would pick out the entire possibility space even before awareness growth. If so, by Awareness Rigidity, $\mathsf{P}^+(tails|heads \vee tails)$ and $\mathsf{P}^+(lion|heads \vee tails)$ should both equal $1/2$ since these were their probabilities before awareness growth. But these assignments would force $\mathsf{P}^+(stonehenge|heads \vee tails)$ to zero. To avoid this odd result for awareness Rigidity, one might argue that $heads \vee tails$ picks out a possibility space larger than the old one, because it also includes the possibility of

These difficulties disappear if the scenario is modeled using Bayesian networks. The definition of awareness expansion we have been working with is simple: whenever a new state is added to one of the nodes in the network, awareness expansion takes place. The novel state can be added to any node in the network. Each node, with its range of states, characterizes an exhaustive partition of the possibility space. Whenever a new state is added to a node, the partition associated with the node expands. (Refinement, as we will see in the next section, consists in the addition of a new node, not in the addition of a new state to an existing node.)

By the definition of expansion just given, COIN counts as a case of expansion, but it is structurally different from the more straightforward cases such as FRIENDS and TENANT. Bayesian networks can help to model the difference precisely. The scenario can be modeled by this familiar graph structure:



The upstream node *outcome* has two states, *tails* and *heads*. These two states remain the same throughout. What changes are the states associated with the *imagine* node downstream. Before awareness growth, the node *image* has two states: *lions* and *heads-image*.[12] You assume that $Image = lions$ is true if and only if $Outcome = tails$ is true. Then, you come to the realization that the imagines for tails could include a lion or a stonehenge engraving. So, after awareness growth, the node *Image* contains three states: *lion*, *stonehenge* and *heads-image*.

To some extent, COIN has the same structure as TENANT—they are modeled by the same networks structure—but there is also an important asymmetry that is apparent by comparing their Bayesian networks. In the network for COIN, the states of the upstream node remain fixed while a new state is added to the downstream node. In the network for TENANT, the opposite happens: the states of the downstream node remain fixed, while a new state is added to the upstream node. Specifically, after awareness expansion, no new state is added to upstream node *Outcome*, but an additional state, *other*, is added to the downstream node *Person*.

The states of the upstream node model what we view—loosely speaking—as more causally fundamental compared to the states of downstream nodes. In TENANT, that a person singing in the shower is more fundamental, while the proposition that describes the person's roles is derivative, for multiple people could play the same role. In *Coin*, that an outcome (heads or

stonehenge. But arguably $heads \vee tails$ should not pick out a larger possibility space. So which is it?

[12] The heads side must have some image, not specified in the scenario.

tails) is instantiated by different engravings is considered a derivative fact. What is causally more fundamental is the outcome, not the engraving. Bayesian networks offer a language to model these differences that are crucial to model episodes of awareness expansion.

So what about constraint (C)? It is easy to check that it is satisfied. Conditional probabilities such as $\mathsf{P}(\mathit{Image} = \mathit{lions}|\mathit{Outcome}= \mathit{tails})$ or $\mathsf{P}(\mathit{Image} = \mathit{lions}|\mathit{Outcome}= \mathit{heads})$ remain unchanged after awareness growth given the condition $\mathit{Image} \neq \mathit{stonehenge}$. Initially, $\mathit{Image} = \mathit{lions}$ is true if and only if $\mathit{Outcome} = \mathit{tails}$ is true. So, $\mathsf{P}(\mathit{Image} = \mathit{lions}|\mathit{Outcome}= \mathit{tails})$ equals one, but it must also be that $\mathsf{P}^{+}(\mathit{Image} = \mathit{lions}|\mathit{Outcome}= \mathit{tails}\ \&\ \mathit{Image} \neq \mathit{stonehenge})$ equals one. More generally, plausible probability distributions for the Bayesian networks associated with the scenario COIN is displayed in Table 2. Constraint (C) is never violated.

All in all, examples in the literature that count as cases of expansion under our definition—that is, a state is added to a network without changes in the network structure—obey constraint (C). This provides good support for the first part of our working hypothesis: if there is no change in the structure of the network, constraint (C) holds generally (see end of Section 3). In this sense, the constraint has outperformed both Reverse Bayesianism and Awareness Rigidity.

But our objective here is not to replace one formal constraint with another. As noted in the introduction, we think that the phenomenon of awareness growth in its generality cannot be modeled in a purely formal matter. The success of constraint (C) relies on the right network structure. How the networks should be built is based on our subject-matter knowledge—for example, that people's behavior must have a reason; that multiple peoples can play different roles; or that heads and tails can be associated with different specific engravings. Constraint (C) holds when this subject-matter knowledge does not change. However, sometimes awareness growth may bring in new subject-matter knowledge and require changes to the structure of the network. This is our next topic.

5 Refinement with Bayesian Networks

To see how the network structure itself may require modifications, we turn now from cases of expansion to cases of refinement. In the framework of Bayesian networks, expansion consists in adding states to existing nodes in the network. Refinement, instead, can be modeled by adding nodes to the network without adding any new state to existing nodes. Intuitively, refinement

$\mathsf{P}(\mathit{Image}\mid\mathit{Outcome})$		*Outcome*	
		heads	*tails*
Image	*lion*	0	1
	heads-image	1	0
	TOTAL	1	1
$\mathsf{P}^{+}(\mathit{Image}\mid\mathit{Outcome})$		*Outcome*	
		heads	*tails*
Image	*lion*	0	1/2
	stonehenge	0	1/2
	heads-image	1	0
	TOTAL	1	1
$\mathsf{P}(\mathit{Outcome}) = \mathsf{P}^{+}(\mathit{Outcome})$	*Outcome*		
	heads	*tails*	
	1/2	1/2	

Table 2: Table displays a plausible probability distribution for the COIN scenario. Constraint (C) is met.

takes place when an epistemic agent acquires a more-fined grained picture of the situation, say instead of thinking that the political spectrum is divided into liberals and conservatives, the political spectrum can be further divided into traditional-liberal, new-liberal, traditional-conservative and new-conservative. The political spectrum is still divided into liberal and conservative—no expansion occurred—but the two categories have been further refined.[13]

Although there is no shortage of counterexamples to Reverse Bayesianism when it comes to awareness refinement, we will use our own. Recall that MOVIES—the refinement-based counterexample to Reverse Bayesianism by Steele & Stefánsson—suffered from a possible objection. The example contained awareness refinement paired with a standard case of Bayesian learning by conditionalization. Some might argue that conditionalization, not awareness refinement, is responsible for the change in probabilities. To alleviate this worry, we will work with our own example which can be more clearly interpreted as mere awareness refinement. Our own example will also allow us to underscore the role of subject-matter assumptions in theorizing about awareness growth. So consider this scenario:

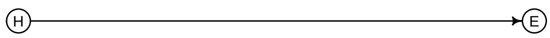
> LIGHTING: You have evidence that favors a certain hypothesis, say a witness saw the defendant around the crime scene. You give some weight to this evidence. In your assessment, that the defendant was seen around the crime scene (your evidence) raises the probability that the defendant was actually there (your hypothesis). But now you ask, what if it was dark when the witness saw the defendant?

[13]This example is from Pettigrew (forthcoming).

> In light of your realization that it could have been dark, you wonder whether (and if so how) you should change the probability that you assigned to the hypothesis that the defendant was around the crime scene.

As your awareness grows, you do not learn anything specific about the lighting conditions, neither that they were bad nor that they were good. You simply wonder what they were, a variable you had previously not considered. So no Bayesian conditoning takes place in the strict sense, although broadly speaking some new information has been introduced.[14] Something has changed in your epistemic state—you have a more fine-grained assessment of what could have happened—but it is not clear what you should do in this scenario. Since the lighting conditions could have been bad but could also have been good, perhaps you should just stay put until you learn something more specific.

In what follows, we illustrate how Bayesian networks helps to model what is going on in LIGHTING and conclude that you should probably revise downward your confidence in the hypothesis that the defendant was around the crime scene. The starting point of our analysis is the usual hypothesis-evidence idiom, repeated below for convenience:



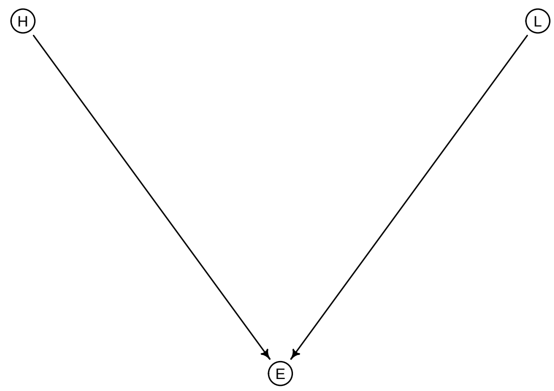
Since you trust the evidence, you think that the evidence is more likely under the hypothesis that the defendant was present at the crime scene than under the alternative hypothesis:

$$\mathsf{P}(E{=}\mathit{seen}|H{=}\mathit{present}) > \mathsf{P}(E{=}\mathit{seen}|H{=}\mathit{absent})$$

The inequality is a qualitative ordering of how plausible the evidence is in light of competing hypotheses. No matter the numbers, by the probability calculus, it follows that the evidence raises the probability of the hypothesis *H=present*.

Now, as you wonder about the lighting conditions, the graph should be amended:

[14]The process of awareness growth in LIGHTING adds only one extra variable, lighting conditions, while MOVIES adds two extra variables, language difficulty and whether the owner is simple-minded or not. Further, MOVIES contains a clear-cut case of Bayesian updating, that the owner *is* simple-minded. This is not so in LIGHTING. Strictly speaking, you are learning that it is *possible* that the lighting conditions were bad. However, you are not conditioning on the proposition 'the lighting conditions were bad' or 'the lighting conditions were good'. So you are not learning about the lighting conditions in the sense of Bayesian updating.



where the node L can have two values, L=*good* and L=*bad*. What is going on here? Initially, you thought that the perceptual experience of the witness (node E) was causally affected by te state of the whereabouts of the defendants (node H). But, as your awareness growth, you reliaze that the witness' experience may also be caused by the environmental condition surrounding the experience itself, say the lighting conditions (node L). So there are now two incoming arrows into node E. In addition, commonsense as well as psychological findings suggest that when the visibility deteriorates, people's ability to identify faces worsen. So a plausible way to modify your assessment of the evidence is as follows:

$$\mathsf{P}^+(E{=}seen|H{=}present \wedge L{=}good) > \mathsf{P}^+(E{=}seen|H{=}absent \wedge L{=}good)$$

$$\mathsf{P}^+(E{=}seen|H{=}present \wedge L{=}bad) = \mathsf{P}^+(E{=}seen|H{=}absent \wedge L{=}bad)$$

In words, if the lighting conditions were good, you still trust the evidence like you did before (first line), but if the lighting conditions were bad, you regard the evidence as no better than chance (second line). These probabilistic constraints are plausible, but should ultimately be grounded on verifiable empirical regularities.

Despite the change in awareness, you have not learned anything in the strict sense. Your new stock of evidence does not contain neither the information that the lighting conditions were bad nor that they were good. But the Bayesian network structure that represents your epistemic state is now more fine-grained. The network contains the new variable L which it did not contain prior to the episode of awareness growth. In addition—and this is the crucial point—the new variable bears certain *structural relationships* with the variables H and E. The graphical network represents a direct probabilistic dependency between the lighting conditions L and the witness sensory experience E, but does not allow for any direct dependency between the lighting conditions and the fact that the defendant was (or was not) at the crime scene. There is no direct arrow between the nodes L and H. This structure of dependencies captures our causal intuitions about the scenario: the lighting conditions do affect what the witness could see, but do not directly affect what the defendant might have done.

Without Bayesian networks, episodes of awareness growth could only be modeled by the addition of new propositions that were not previously given in the algebra. But this approach fails to capture crucial information. When awareness growth takes place against the background of an intuitive causal structure of the world—as in the case of LIGHTING—this structure should also be modeled. Bayesian networks offer a formal framework that can do precisely that.

This model of the underlying causal structure can now guide us to decide whether our revised version of Reverse Bayesianism, what we called constraint (C), holds in this scenario. Specifically, we need to assess whether the following holds:

$$\frac{\mathsf{P}(E = seen|H = present)}{\mathsf{P}(E = seen|H = absent)} = \frac{\mathsf{P}^{+}(E = seen|H = present)}{\mathsf{P}^{+}(E = seen|H = absent)}.$$

The question here is whether you should assess the evidence at your disposal—that the witness saw the defendant at the crime scene—any differently than before.[15] As noted earlier, without a clear model of the scenario, it might seem that you should simply stay put. After all, besides the sensory experience of the witness, you have gained no novel information about the lighting conditions. Should you thus conclude that the evidence has the same value before and after the realization that lighting could have been bad?

The evidence would have the same value if the likelihood ratios associated with it relative to the competing hypotheses were the same before and after awareness growth. But, in changing the probability function from $\mathsf{P}()$ to $\mathsf{P}^{+}()$, it would be quite a coincidence if this were true. In our example, many possible probability assignments violate this equality. If before awareness growth you thought the evidence favored the hypothesis H=*present* to some extent, after the growth in awareness, the evidence is likely to appear less strong. To see this is tedious, so we relegate the details to a footnote.[16] If this is correct, constraint (C) has been violated. Reverse

[15]Note that since no new state was added to an existing node, the condition $X \neq x^*$ in constraint (C) (where x^* is the new state added to an existing node X) is redundant here.

[16]By the law of total probability, the right hand side of the equality in (C) should be expanded, as follows:

$$\frac{\mathsf{P}^{+}(E = e|H = h)}{\mathsf{P}^{+}(E = e|H = h')} = \frac{\mathsf{P}^{+}(E{=}seen \wedge L{=}good|H{=}present) + \mathsf{P}^{+}(E{=}seen \wedge L{=}bad|H{=}present)}{\mathsf{P}^{+}(E{=}seen \wedge L{=}good|H{=}absent) + \mathsf{P}^{+}(E{=}seen \wedge L{=}bad|H{=}absent)}.$$

Bayesianism, in its orginal formulation, is also violated since the ratio of the probabilities of *H=present* to *E=seen*, before and after awareness growth, has changed:

$$\frac{\mathsf{P}^{E=seen}(H{=}present)}{\mathsf{P}^{E=seen}(E{=}seen)} \neq \frac{\mathsf{P}^{+,E=seen}(H{=}present)}{\mathsf{P}^{+,E=seen}(E{=}seen)},$$

where $\mathsf{P}^{E=seen}()$ and $\mathsf{P}^{+,E=seen}()$ represent the agent's degrees of belief, before and after awareness growth, updated by the evidence *E=seen*.[17]

The general lesson to be learned here has to do with the importance of formalizing subject-matter assumptions and the role of Bayesian networks in modeling awareness growth. Modeling the causal assumptions in LIGHTING allowed us to see that constraint (C)—as well as Reverse Bayesianism more generally—should fail here.

On the other hand, constraint (C) holds in other, structurally different cases of refinement.

For concreteness, let's use some numbers:

$$\mathsf{P}(E{=}seen|H{=}present) = \mathsf{P}^{+}(E{=}seen|H{=}present \wedge L{=}good) = .8$$

$$\mathsf{P}(E{=}seen|H{=}absent) = \mathsf{P}^{+}(E{=}seen|H{=}absent \wedge L{=}good) = .4$$

$$\mathsf{P}^{+}(E{=}seen|H{=}present \wedge L{=}bad) = \mathsf{P}^{+}(E{=}seen|H{=}absent \wedge L{=}bad) = .5.$$

$$\mathsf{P}^{+}(L{=}bad) = \mathsf{P}^{+}(L{=}good) = .5.$$

So the ratio $\frac{\mathsf{P}(E=seen|H=present)}{\mathsf{P}(E=seen|H=absent)}$ equals 2. After the growth in awareness, the ratio $\frac{\mathsf{P}^{+}(E=seen|H=present)}{\mathsf{P}^{+}(E=seen|H=absent)}$ will drop to $\frac{.65}{.45} \approx 1.44$. The calculations here rely on the dependency structure encoded in the Bayesian network (see starred step below).

$$\begin{aligned}
\mathsf{P}^{+}(E{=}seen|H{=}present) &= \mathsf{P}^{+}(E{=}seen \wedge L{=}good|H{=}present) + \mathsf{P}^{+}(E{=}seen \wedge L{=}bad|H{=}present)\\
&= \mathsf{P}^{+}(E{=}seen|H{=}present \wedge L{=}good) \times \mathsf{P}^{+}(L{=}good|H{=}present)\\
&+ \mathsf{P}^{+}(E{=}seen|H{=}present \wedge L{=}bad) \times \mathsf{P}^{+}(L{=}bad|H{=}present)\\
&=^{*} \mathsf{P}^{+}(E{=}seen|H{=}present \wedge L{=}good) \times \mathsf{P}^{+}(L{=}good)\\
&+ \mathsf{P}^{+}(E{=}seen|H{=}present \wedge L{=}bad) \times \mathsf{P}^{+}(L{=}bad)\\
&= .8 \times .5 + .5 * .5 = .65
\end{aligned}$$

$$\begin{aligned}
\mathsf{P}^{+}(E{=}seen|H{=}absent) &= \mathsf{P}^{+}(E{=}seen \wedge L{=}good|H{=}absent) + \mathsf{P}^{+}(E{=}seen \wedge L{=}bad|H{=}absent)\\
&= \mathsf{P}^{+}(E{=}seen|H{=}absent \wedge L{=}good) \times \mathsf{P}^{+}(L{=}good|H{=}absent)\\
&+ \mathsf{P}^{+}(E{=}seen|H{=}absent \wedge L{=}bad) \times \mathsf{P}^{+}(L{=}bad|H{=}absent)\\
&=^{*} \mathsf{P}^{+}(E{=}seen|H{=}absent \wedge L{=}good) \times \mathsf{P}^{+}(L{=}good)\\
&+ \mathsf{P}^{+}(E{=}seen|H{=}absent \wedge L{=}bad) \times \mathsf{P}^{+}(L{=}bad)\\
&= .4 \times .5 + .5 * .5 = .45
\end{aligned}$$

This argument can be repeated with many other numerical assignments.

[17] The scenario also violates Awareness Rigidity which requires that $\mathsf{P}^{+}(A|T^{*}) = \mathsf{P}(A)$, where T^{*} corresponds to a proposition that picks out, from the vantage point of the new awareness state, the entire possibility space before the episode of awareness growth. In LIGHTING, however, T^{*} does not change, so Awareness Rigidity would require that $\mathsf{P}^{+}(A) = \mathsf{P}(A)$, and instead in the scenario, we have

$$\mathsf{P}^{+}(H{=}present|E{=}seen) \neq \mathsf{P}(H{=}present|E{=}seen).$$

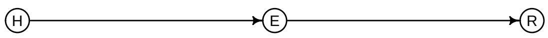
For consider this scenario:

> VERACITY: A witness saw that the defendant was around the crime scene and you initially took this to be evidence that the defendant was actually there. But then you worry that the witness might be lying or misremembering what happened. Perhaps, the witness was never there, made things up or mixed things up. Should you reassess the evidence at your disposal? If so, how?

This scenario might seem no different from LIGHTING. The realization that lighting could be bad should make you less confident in the truthfulness of the sensory evidence. And the same conclusion should presumably follow from the realization that the witness could be lying. But, upon closer scrutiny, running the two scenarios together turns out to be a mistake.

The evidence at your disposal in LIGHTING is the sensory evidence—the experience of seeing—and the possibility of bad lighting does affect the quality of your visual experience. So, if lighting was indeed bad, this would warrant lowering your confidence in the truthfulness of the visual experience. But the possibility of lying in VERACITY does not affect the quality of the visual experience; it only affects the quality of the *reporting* of that experience. So, if the witness did lie, this would not warrant lowering your confidence in the truthfulness of the visual experience, only in the truthfulness of the reporting. The distinction between the visual experience and its reporting is crucial here. Bayesian networks help to model this distinction precisely, and then see why LIGHTING and VERACITY are structurally different.

The graphical network should initially look like the initial DAG for LIGHTING, consisting of the hypothesis node H upstream and the evidence node E downstream. As your awareness grows, the graphical network should be updated by adding another node R further downstream:



As before, the hypothesis node H bears on the whereabouts of the defendant and has two values, H=*present* and H=*absent*. Note the difference between E and R. The evidence node E bears on the visual experience had by the witness. The reporting node R, instead, bears on what the witness reports to have seen. The chain of transmission from 'visual experience' to 'reporting' may fail for various reasons, such as lying or misremembering.

In VERACITY, the conditional probabilities, $\mathsf{P}(E=e|H=h)$ should be the same as $\mathsf{P}^+(E=e|H=h)$ for any values e and h of the variables H and E that are shared before and after awareness growth. In comparing the old and new Bayesian network, this equality falls out from their structure, as the connection between H and E remains unchanged. Thus, constraint (C) is perfectly fine in scenarios such as VERACITY.

This does not mean that the assessment of the probability of the hypothesis *H=present* should undergo no change. If you worry that the witness could have lied, this should presumably make you less confident about *H=present*. To accommodate this intuition, VERACITY can be interpreted as a scenario in which an episode of awareness refinement takes place together with a form of retraction. At first, after the learning episode, you update your belief based on the *visual experience* of the witness. But after the growth in awareness, you realize that your learning is in fact limited to what the witness *reported* to have seen. The previous learning episode is retracted and replaced by a more careful statement of what you learned: instead of conditioning on *E=seen*, you should condition on what the witness reported to have seen, *R=seen-reported*. This retraction will affect the probability of the hypothesis *H=present*.

Where does this leave us? Refinement cases that might at first appear similar can be structurally different in important ways, and this difference can be appreciated by looking at the Bayesian networks used to model them. In modeling VERACITY, the new node is added downstream, while in modeling LIGHTING, it is added upstream. This difference affects how probability assignments should be revised. Since the conditional probabilities associated with the upstream nodes are unaffected, constraint (C) is satisfied in VERACITY. By contrast, since the conditional probabilities associated with the downstream node will often have to change, the constraint fails in LIGHTING. This discussion further corroborates the claim that subject-matter assumptions—in the case of LIGHTING and VERACITY, causal assumptions—about how we conceptualize a specific scenario are the guiding principles for how we should update the probability function through awareness growth, not formal principles like Reverse Bayesianism, Awareness Rigidity or even constraint (C).

6 Towards a general theory

We conclude by sketching a recipe of how to model awareness growth using Bayesian network. Awareness growth can occur in different ways. The key question is to what extent probability assignments that were made prior to the episode of awareness growth can be retained. There seems to no clear rule to decide that. The problem of awareness growth can be formulated as follows: when new propositions are added to the algebra, to what extent can the old probabilities assignments be preserved? To answer this question, we propose the following procedure.

The first step is to draw a direct acyclical graph $\mathscr{G}$ that expresses the probabilistic dependencies between the variables (and their relative states) *before* awareness growth. This graphical structure represents the material structural assumptions, based on commonsense, semantic stipulations or causal dependency.[18] The graphical structure should be accompanied by probability tables. The graph and the probability tables define the full probability distribution $\mathsf{P}()$.

The next step is to decide what changes to the graphical structure $\mathscr{G}$ must be made to adequately model awareness growth. In general, awareness growth will require either to add states to the existing nodes (expansion) or to modify the structure of the network (refinement). So we should consider each case in turn.

Suppose awareness growth is modeled by adding a state x^* to an existing node X in the network. Then, the existing probability tables in which X occurs should be modified to accommodate the new value x^* of X. Specifically, the probability table for X should be modified, as well as the probability tables for the children nodes of X. Constraint (C) will guide on how to change the probability tables in question and define the new probability distribution $\mathsf{P}^+()$. No other changes to the probability tables are necessary because the theory of Bayesian networks ensure that a variable X is independent from all other variables conditional on its so-called Markov blanket (which includes the parent and children nodes of X, as well as the parents of the children of X).

Suppose instead awareness growth is modeled by adding a new node X to the network. There are different cases to distinguish here. In the easiest case, the new node is added without breaking existing connections between nodes. The addition of the new node only requires drawing additional arrows that connect the new node to existing nodes. Both VERACITY and LIGHTING are such cases. In these two scenarios, the new nodes X is either added downstream of an existing node or downstream. A slightly more complicated case is one in which the new node X is both downstream (relative to an existing node A) and upstream (relative to another existing node B).[19] Another option is that the new node is downstream relative to both

[18] Arrows in Bayesian networks are often taken to represent causal relationships, but other interpretations exist. Schaffer (2016) discusses an interpretation in which arrows represent grounding relations rather than causality.

[19] For example, initially you thought that gender (G) had an effect on graduate admission decision (D). The aggregate data available indicate that women are less likely to be admitted to graduate schools. So the initial

existing nodes or upstream relative to both. In other words, the new node creates via the novel connections an alternative path between the two variables A and B. Here is a summary of the cases:

$\mathscr{G}$	A	$\rightarrow$	B
$\mathscr{G}^{+}$		X	
(X is upstream, e.g. LIGHTING)	$\swarrow$	X	
	A	$\rightarrow$	B
(X is downstream, e.g. VERACITY)		X	$\nwarrow$
	A	$\rightarrow$	B
(X is common cause)	$\swarrow$	X	$\searrow$
	A	$\rightarrow$	B
(X is common effect)	$\nearrow$	X	$\nwarrow$
	A	$\rightarrow$	B

Then, a new probability table should be added, specifically, a new probability table for X. In addition, existing probability tables should be modified, should be added and existing ones should be modified (specifically, the probability tables of the nodes that are now the children and te parents of X should be modified). See the Appendix for details.

Besides the changes in the previous steps, no other changes will be required to the probabilityu tables of the new Bayesian network to define the new probability distribution $\mathsf{P}^{+}()$.

We emphasis that this procedure is only partly algorithmic. A key role play the background information that is formalized in the structure of the Bayesian network before and after awareness growth. This procedure implements a conservative constraint by which minimal changes are made to the old probability distribution. For the changes are localized to the probability tables of the node X and the children of node X. These localized changes are justified on the basis of the dependencies structure among nodes of the old and new Bayesian network. If the modification of to an existing nodes or the addition of an existing node does require changes to the conditional probabilities of other nodes—except the node itself and its children—this means that the changes are irrelevant to the other nodes.

graph looks like this: $G \rightarrow D$. You then hypothesize that, since schools (S) make decision about admission, not the university as a whole, there might an alternative path from G to D. Perhaps, women happen to prefer departments that have lower admission rates. So a new path must be added to the graph: $G \rightarrow S \rightarrow D$.

7 Appendix

7.1 Refinement: adding states to existign nodes

Suppose X is the node to which a new state x^* must be added. Let $\mathbf{U}(X)$ be set of immediate upstream nodes of X (the parents of X) and let $\mathbf{D}(X)$ be set of immediate downstream nodes of X (the children of X). So we have:

$$\mathbf{U}(X) \rightrightarrows X \rightrightarrows \mathbf{D}(X),$$

where the symbol $\rightrightarrows$ indicates there could be multiple arrows from nodes in $\mathbf{U}(X)$ into X and multiple arrows from X into nodes in $\mathbf{D}(X)$. Before the addition of the new state x^*, the old probability distribution was defined by probability tables for these conditional probabilities:

- $\mathsf{P}(X = x | U_1 \wedge \cdots \wedge U_n)$, for all values of x of X drawn from the set $\{x_1, \ldots, x_k\}$ and all permutations of values of the upstream nodes $U_1, \ldots, U_n$ in $\mathbf{U}(X)$.
- $\mathsf{P}(D_i | X = x)$, for all values of x of X drawn from the set $\{x_1, \ldots, x_k\}$ and all values of any downstream node D_i in $\mathbf{D}(X)$.

When the state x^* is added to the values of X, the values x of X should now be drawn from the extended set $\{x_1, \ldots, x_k, x^*\}$. So we have:

- $\mathsf{P}^+(X = x | U_1 \wedge \cdots \wedge U_n)$, for all values of x of X drawn from the set $\{x_1, \ldots, x_k, x^*\}$ and all permutations of values of the upstream nodes $U_1, \ldots, U_n$ in $\mathbf{U}(X)$.
- $\mathsf{P}^+(D_i | X = x)$, for all values of x of X drawn from the set $\{x_1, \ldots, x_k, x^*\}$ and all values of any downstream node D_i in $\mathbf{D}(X)$.

A generalized version of constraint (C) governs the relationship between $\mathsf{P}()$ and $\mathsf{P}^+()$, as follows:

- $\mathsf{P}(X = x | U_1 \wedge \cdots \wedge U_n) = \mathsf{P}^+(X = x | U_1 \wedge \cdots \wedge U_n \wedge X \neq x^*)$, for all values of x of X and all permutations of values of the upstream nodes $U_1, \ldots, U_n$ in $\mathbf{U}(X)$.
- $\mathsf{P}(D_i | X = x) = \mathsf{P}^+(D_i | X = x \wedge X \neq x^*)$, for all values of x of X and all values of any downstream node D_i in $\mathbf{D}(X)$.

7.2 Expansion: adding nodes and arrows

Suppose X is the new node to be added to the network, with $\mathbf{U}(X)$ the immediate upstream nodes of X (the parents of X) and $\mathbf{D}(X)$ the immediate downstream nodes of X (the children of X). Relative to the upstream nodes to X, new probability tables should be defined for the following conditional probabilities:

- $\mathsf{P}^+(X = x|U_1 \wedge \cdots \wedge U_n)$, for all values of x of X and all permutations of values of the upstream nodes $U_1, \ldots, U_n$ in $\mathbf{U}(X)$.

If the new node X does not have any downstream node, there will be no further changes needed. This was the case with the scenario VERACITY.

But suppose the new node X has also downstream nodes, or in other words it is placed upstream relative to other nodes. Node X could have multiple downstream nodes. Let $D \in \mathbf{D}(X)$ be one such node and let $\mathbf{U}(D)$ the set of immediate upstream nodes of D (the parents of D). Then, the new probability table should be defined for the following conditional probabilities:

- $\mathsf{P}^+(D = d|X = x \wedge U_1 \wedge \cdots \wedge U_n)$, for all values of d of $D \in \mathbf{D}(X)$ and x of X and all permutations of values of the upstream nodes $U_1, \ldots, U_n$ in $\mathbf{U}(D)$ (that is, the upstream nodes of D which is one of the downstream nodes of X.)
- In contrast, the old conditional probabilities were instead defined as $\mathsf{P}(D = d|U_1 \wedge \cdots \wedge U_n)$, for all values of d of D and all permutations of values of the upstream nodes $U_1, \ldots, U_1$ in $\mathbf{U}(D)$. So the statement $X = x$ was added.

The same applies to all the other downstream nodes D of X, that is to all node D's in $\mathbf{D}(X)$.

References

Bradley, R. (2017). *Decision theory with a human face*. Cambridge University Press.

Chihara, C. S. (1987). Some problems for bayesian confirmation theory. *British Journal for the Philosophy of Science*, *38*(4), 551–560.

Earman, J. (1992). *Bayes or bust? A critical examination of bayesian confirmation theory*. MIT press.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A General Structure for Legal Arguments About Evidence Using Bayesian Networks. *Cognitive Science*, *37*(1), 61–102. https://doi.org/10.1111/cogs.12004

Glymour, C. (1980). *Theory and evidence*. Princeton University Press.

Howson, C. (1976). The development of logical probability. In *Essays in memory of imre lakatos. Boston studies in the philosophy of science* (pp. 277–298). Springer.

Karni, E., & Vierø, M.-L. (2015). Probabilistic sophistication and reverse bayesianism. *Journal of Risk and Uncertainty Volume*, *50*, 189–208.

Lakatos, I. (1968). Changes in the problem of inductive logic. *Studies in Logic and the Foundations of Mathematics*, *51*, 315–417.

Mathani, A. (2020). Awareness growth and dispositional attitudes. *Synthese*, *198*(9), 8981–8997.

Pettigrew, R. (forthcoming). How should your beliefs change when your awareness grows? *Episteme*, 1–25. https://doi.org/10.1017/epi.2022.33

Roussos, J. (2021). Awareness growth and belief revision. *Manuscript*.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies*, *173*, 49–100.

Steele, K., & Stefánsson, O. (2021). Belief revision for growing awareness. *Mind*, *130*(520), 1207–1232.

Wenmackers, S., & Romeijn, J.-W. (2016). New theory about old evidence: A framework for open-minded bayesianism. *Synthese Volume*, *193*, 1225–1250.

Williamson, J. (2003). Bayesianism and language change. *Journal of Logic, Language, and Information*, *12*(1), 53–97.

Zabell, S. (1992). Predicting the unpredictable. *Synthese*, *90*(1), 205–232.