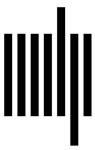
The Handbook of Rationality

Edited by: Markus Knauff, Wolfgang Spohn

Citation:

The Handbook of Rationality
Edited by: Markus Knauff, Wolfgang Spohn
DOI: 10.7551/mitpress/11252.001.0001
ISBN (electronic): 9780262366175
Publisher: The MIT Press
Published: 2021

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.



The MIT Press

4.2 Bayes Nets and Rationality

Stephan Hartmann

Summary

Bayes nets are a powerful tool for researchers in statistics and artificial intelligence. This chapter demonstrates that they are also of much use for philosophers and psychologists interested in (Bayesian) rationality. To do so, we outline the general methodology of Bayes nets modeling in rationality research and illustrate it with several examples from the philosophy and psychology of reasoning and argumentation. Along the way, we discuss the normative foundations of Bayes nets modeling and address some of the methodological problems it raises.

1. The Virtues of Bayes Nets

Bayes nets (or "Bayesian networks") are a powerful tool for researchers in statistics and artificial intelligence. The reason for this consists in the many scientific virtues of Bayes nets. One of these virtues is their *representational power*: Bayes nets represent in an intuitive way the conditional independencies that hold between variables. Exploiting these conditional independencies, Bayes nets allow for a compact representation of a joint probability distribution over n variables: what would ordinarily, for binary variables without any additional constraints, require the specification of $2^n - 1$ parameters here needs significantly less. This has enormous practical advantages and hence initiated the development of probabilistic expert systems and other technical applications. Another virtue of Bayes nets is their *algorithmic power*: Bayes nets come with efficient algorithms to derive from the joint distribution whatever marginal or conditional probability one is interested in. Last but not least, the theory of Bayes nets is very elegant, and not much is needed to successfully apply it to new problems.

These are some of the reasons why Bayes nets have already found so many applications in various parts of science and engineering. The goal of this chapter is to demonstrate that Bayes nets are also of much use for philosophers and psychologists in the field of (Bayesian) rationality. We will see that Bayes nets can be naturally integrated into the Bayesian framework and help solving problems that would otherwise be hard to address. In the following pages, we will outline the general methodology of Bayes nets modeling in rationality research and elaborate on the various functions of these models. The methodology will then be illustrated by analyzing a number of problems and questions from the philosophy and psychology of reasoning and argumentation.

The remainder of this chapter is organized as follows: section 2 provides a concise introduction to the theory of Bayes nets. Section 3 introduces Bayesian rationality. Here we distinguish between the general Bayesian framework and the models that are constructed within the framework to address a specific problem or question from rationality research. Bayes nets play a role in the latter but not in the former. We illustrate the methodology by examining a number of increasingly complex confirmation scenarios. Section 4 considers two examples from the philosophy and psychology of reasoning and argumentation in more detail. We will see that Bayes nets models help the rationality researcher to reconstruct certain reasoning and argumentation schemes and to identify possible holes in an argument (and to suggest a remedy). The section closes with the sketch of a general theory of Bayesian argumentation with a special focus on the role of indicative conditionals. It builds heavily on the use of Bayes nets. Section 5, finally, closes with a short outlook.

2. Bayes Nets in a Nutshell

A Bayes net organizes a set of variables into a *directed acyclic graph* (DAG). A DAG is a set of nodes and a set of arrows between those nodes. The only constraint is that there are no closed paths formed by following the arrows. A "root node" is a node with outgoing arrows only, a "parent" of a given node is a node from which an arrow points at the given node, and a "descendant" of a node is one that is pointed at by a corresponding

arrow. Each node represents a propositional variable, which can take any number of mutually exclusive and jointly exhaustive values. To make a DAG into a Bayes net, one more step is required: we need to specify the prior probabilities for the variables in the root nodes and the conditional probabilities for the variables in all other nodes, given any combination of values of the variables in their respective parent nodes.

The arrows in a Bayes net carry information about the conditional independence relations between the variables in the DAG. This information is expressed by the *Parental Markov Condition* (PMC):

PMC A variable represented by a node in a Bayes net is independent of all variables represented by its non-descendant nodes, conditional on all variables represented by its parents.

Here is an illustration: consider the Bayes net in figure 4.2.1, involving the three propositional variables A, B, and C. Node C is a root node. It is the parent of A and B, and A and B are the children of C. Applying PMC, we find that $A \perp\!\!\!\perp B \mid C$ (read: A is independent of B given C). One also says that C *screens off* A from B.

Compare this Bayes net with the one in figure 4.2.2. Here B is the parent of C and C is the parent of A. Node A is a child of C and at the same time a descendant of B. Interestingly, applying PMC, we find that $A \perp\!\!\!\perp B \mid C$ also. The networks in figures 4.2.1 and 4.2.2 therefore represent the same conditional independence structure. That is, if a modeler has reason to assume that $A \perp\!\!\!\perp B \mid C$, then this conditional independence can be represented in a Bayes net in two ways. To single out one of them, further information is needed.

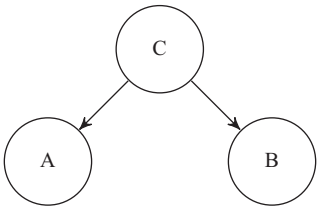


Figure 4.2.1
The "common cause" network.

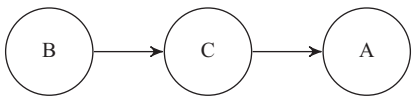


Figure 4.2.2
The chain network.

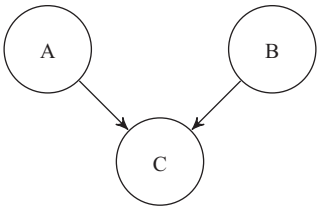


Figure 4.2.3
The collider (or "common effect") network.

Finally, let us consider the Bayes net in figure 4.2.3. Here the variables A and B are the parents of C. Applying PMC, we find that $A \perp\!\!\!\perp B$, that is, A is unconditionally independent of B (one could also write $A \perp\!\!\!\perp B \mid \emptyset$, where $\emptyset$ represents the empty set). Hence, the conditional independence structure instantiated in the Bayes net in figure 4.2.3 differs from the one in figures 4.2.1 and 4.2.2.

Conditional independence structures can be investigated in general, and there is a rich literature on the topic. Most important, conditional independence structures (i.e., the three-place relation $\cdot \perp\!\!\!\perp \cdot \mid \cdot$) satisfy the so-called *semi-graphoid axioms* (A. P. Dawid, 1979; Spohn, 1980), which can be used to derive new conditional independencies from already known ones.[1] This is important as it is easy to see that PMC does not allow us to identify *all* conditional independencies that hold in a DAG. Consider the Bayes net in figure 4.2.2. Here PMC implies only that $A \perp\!\!\!\perp B \mid C$. However, it also seems to be the case that $B \perp\!\!\!\perp A \mid C$, which does not follow from PMC. This conditional independence follows from $A \perp\!\!\!\perp B \mid C$ and the *symmetry axiom* (which is one of the semi-graphoid axioms). Applying the semi-graphoid axioms to find all conditional independencies is rather cumbersome, and so it would be helpful to have one criterion that identifies all conditional independencies in a DAG in a straightforward way. This is the *d*-separation criterion, which is explained in textbooks such as Darwiche (2014) and Pearl (1988).

While conditional independence structures can be studied abstractly, we are only interested in *probabilistic* conditional independence structures. For these, it is helpful to recall that two propositional variables A and B are probabilistically independent with respect to a probability measure P if and only if $P(A,B) = P(A)P(B)$ for all values of A and B. Equivalently, A and B are probabilistically independent with respect to a probability measure P if and only if $P(A \mid B) = P(A)$ for all values of A and B, where $P(A \mid B)$ stands for the conditional probability of A given B, which is defined as $P(A,B)/P(\mathrm{B})$ if $P(\mathrm{B}) > 0$.[2] Generalizing this definition, two propositional variables A and B are probabilistically independent given the

variable C with respect to a probability measure P if and only if $P(A,B\,|\,C) = P(A\,|\,C)P(B\,|\,C)$ for all values of A, B, and C. It is easy to see that this definition is equivalent to the following one: A and B are probabilistically independent given C with respect to a probability measure P iff $P(A\,|\,B,C) = P(A\,|\,C)$ for all values of A, B, and C. That is, once the value of C is known, learning the value of B does not change the probability of A (provided that all conditional probabilities are defined).

While conditional independencies can be read off from the network structure, a probability distribution defined over the DAG is needed to make specific probabilistic inferences. To do so, we express the joint probability distribution over a set of variables $A_1, \ldots, A_n$, which is organized into a DAG in terms of the prior probabilities of all root nodes and the conditional probabilities of all child nodes, given any combination of values of the variables in their respective parent nodes. Let $\mathrm{pa}(A_i)$ denote the set of parents of A_i. Then an application of the chain rule of the probability calculus yields the following expression ("the product rule") for the joint probability distribution over all variables:[3]

$$\begin{aligned} P(A_1, \ldots, A_n) &= \prod_{i=1}^{n} P(A_i\,|\,\mathrm{pa}(A_i)) \\ &= P(A_1\,|\,\mathrm{pa}(A_1)) \cdot P(A_2\,|\,\mathrm{pa}(A_2)) \cdot \cdots \cdot P(A_n\,|\,\mathrm{pa}(A_n)). \end{aligned} \tag{1}$$

Let us illustrate the application of equation 1 with the DAG in figure 4.2.3. To make the DAG a Bayes net, we assume that all variables are binary (with their values represented by A, ¬A, etc.) and specify the (unconditional) probabilities of all root nodes, that is, $P(\mathrm{A}) = a$ and $P(\mathrm{B}) = b$, and the conditional probabilities of the child node given the four combinations of values of the two variables in its respective parent nodes, that is,

$$\begin{aligned} P(\mathrm{C}\,|\,\mathrm{A},\mathrm{B}) &= \alpha, & P(\mathrm{C}\,|\,\mathrm{A},\neg\mathrm{B}) &= \beta, \\ P(\mathrm{C}\,|\,\neg\mathrm{A},\mathrm{B}) &= \gamma, & P(\mathrm{C}\,|\,\neg\mathrm{A},\neg\mathrm{B}) &= \delta. \end{aligned}$$

The product rule (i.e., equation 1) then allows us to compute whatever marginal or conditional probability we are interested in. For example, one easily sees that $P(\mathrm{A},\mathrm{B},\mathrm{C}) = P(\mathrm{A})P(\mathrm{B})P(\mathrm{C}\,|\,\mathrm{A},\mathrm{B}) = ab\alpha$ and that $P(A, \neg B, C) = P(A)P(\neg B)P(C\,|\,A,\neg B) = a\bar{b}\beta$, where we have used the shorthand notation $\bar{x} := 1 - x$, which we will also use below.

Similarly, one can calculate $P(\mathrm{A}\,|\,\mathrm{B},\mathrm{C})$ and $P(\mathrm{A}\,|\,\mathrm{C})$ and show that the two probabilities are *not* identical. That is, A and B are unconditionally independent but conditionally dependent, given C. Fixing the value of the "common effect" variable C renders the "causes" A and B dependent.[4]

3. Bayesian Rationality

Bayesianism is the leading theory of uncertain reasoning.[5] Its starting point is the psychological truism that people believe contingent propositions such as "It will rain tomorrow" more or less strongly: they assign a certain *degree of belief* to a proposition. But what are rational degrees of belief? How can they be combined? And how should one change them if new evidence becomes available? To address these questions, we need a calculus for the representation of degrees of belief (i.e., a theory about the statics of rational belief), rules for changing them (i.e., a theory about the dynamics of rational belief), and a normative foundation for both.

Before moving on, it is useful to distinguish between the *Bayesian framework* and the *models* that are constructed within this framework. While the framework lays out the general features of Bayesian rationality and comes with a *normative foundation*, the models represent specific reasoning situations and help the researcher to tackle concrete problems. These models often involve Bayes nets.

3.1 The Bayesian Framework

Let us begin with the *static part* of Bayesianism. Here Bayesians identify degrees of belief with probabilities. As a consequence, the probability calculus puts specific constraints on the degrees of belief of an agent. For instance, if a rational agent assigns a degree of belief of .3 to the proposition "It will rain tomorrow," then this agent has to assign a degree of belief of .7 to the proposition "It will not rain tomorrow," as the latter proposition is the logical negation of the former and the probability of a proposition and its negation sum up to 1. More generally, a probability distribution P is defined over a Boolean algebra $\mathscr{B}$ of propositions, which comes with rules for the combination (∧ and ∨) and negation (¬) of propositions. In a first step, the agent fixes the algebra, that is, she identifies all relevant propositions.[6] In the second step, the agent specifies a joint probability distribution over the algebra $\mathscr{B}$. As a result, the beliefs of the agent are *coherent*.

Turning to the *dynamic part*, Bayesians specify rules for changing ("updating") probabilities once new information becomes available. Let us assume, for example, that an agent has partial beliefs about the propositions A, B, and C. They are represented by propositional variables A, B, and C, and a prior probability distribution P is defined over them. The agent then learns that A is the

case. That is, the new probability of A, that is, $P'(\mathrm{A})$, is 1. Here P' denotes the new ("posterior") probability distribution of the agent. So far, we only know the new value of the probability of A. But what are, for example, the new probabilities of B and C? And what is the full new joint probability distribution over A, B, and C? Bayesians argue that in this case, the agent should update her probability distribution according to the principle of Conditionalization ("Bayes' theorem"), that is, she should set

$$P'(\mathrm{X}) = P(\mathrm{X} \mid \mathrm{A})$$

for any proposition X in the algebra $\mathscr{B}$ under consideration.

But why should we identify degrees of belief with probabilities? And why should one update according to Conditionalization? That is, what is the *normative foundation* of the Bayesian framework? There are different ways to provide such a normative foundation, and it is a strength of Bayesianism that it is supported by a wide variety of such arguments. The two most popular types of arguments are pragmatic ("Dutch book arguments") and epistemic ("epistemic utility theory"). They are explained in chapter 4.1 by Hájek and Staffel (this handbook).

It is interesting to note that the dynamic part of Bayesianism (i.e., Conditionalization) can also be justified in a different way. To do so, we assume that the agent wants to be as conservative as possible with regard to changing her beliefs. That is, the agent is undogmatic and willing to modify her beliefs once new information comes in, but she wants to make sure that overall, her beliefs change as little as possible. This seems to be psychologically plausible and is also part of other theories of belief revision such as the AGM model, which is explained in chapter 5.2 by Rott (this handbook). More specifically, the principle of *Conservativity* demands that an agent who learns a new item of information make sure that the new probability distribution P' takes this new information into account as a constraint but also requests that P' differ as little as possible from the old (prior) probability distribution P.

Making this idea precise requires the specification of a measure of the distance between two probability distributions. It turns out that the most interesting and useful measures do not satisfy the axioms of a mathematical distance (i.e., of a metric space). For instance, the Kullback–Leibler divergence is not symmetrical and violates the triangle inequality. However, minimizing the Kullback–Leibler divergence yields Conditionalization if one takes into account the constraint that the probability of some proposition in the algebra shifts to 1 (see Diaconis & Zabell, 1982; Eva, Hartmann, & Rafiee Rad, 2020).

It is instructive to note an analogy to Newtonian mechanics here. Newtonian mechanics, too, provides a modeling framework. It specifies a static part (mass points, etc.) and a dynamic part (Newton's second law as a general dynamical law). Furthermore, there are justifications for both parts. What is more, the Newtonian framework comes with various assessment criteria and a (perhaps somewhat implicit) methodology for model construction (see Giere, 1990). This also holds for Bayesian modeling, which is the topic of the following subsection. Before that, however, we introduce Bayesian Confirmation Theory (BCT), which is the central philosophical application of Bayesianism. Its most general aspects are part of the Bayesian framework.

While qualitative confirmation theories formulate criteria that inform us whether or not a piece of evidence E confirms a hypothesis H (Sprenger, 2011), quantitative theories of confirmation (such as BCT) also tell us *how much* E confirms H. According to BCT, an agent starts with a subjective degree of belief that a certain hypothesis H is true—the *prior probability* $P(\mathrm{H})$ of the hypothesis. In the next step, a more or less expected piece of evidence, E, comes in. While E was uncertain before, it now becomes certain. To make sure that the beliefs of the agent remain coherent, the agent updates the probability of H and assigns a *posterior probability* $P'(\mathrm{H})$ to the hypothesis using Conditionalization, that is, $P'(\mathrm{H}) = P(\mathrm{H} \mid \mathrm{E})$. This can also be expressed as

$$P'(\mathrm{H}) = \frac{P(\mathrm{H})}{P(\mathrm{H}) + P(\neg \mathrm{H})x},$$

with the *likelihoods* $p := P(\mathrm{E} \mid \mathrm{H})$ and $q := P(\mathrm{E} \mid \neg \mathrm{H})$ and the *likelihood ratio* $x := q/p$. Now E *confirms* H iff the posterior probability $P'(\mathrm{H})$ (after learning E) is greater than the prior probability $P(\mathrm{H})$. Evidence E *disconfirms* (or "falsifies") H iff the posterior is smaller than the prior. If $P'(\mathrm{H}) = P(\mathrm{H})$, then E is *irrelevant* for H. Equivalently, E confirms H iff the likelihood ratio $x < 1$, E disconfirms H iff $x > 1$, and E is irrelevant for H iff $x = 1$. For more on BCT, see Crupi (2016), Huber (2007), and chapter 4.3 by Merin (this handbook), which also discusses how to measure evidential relevance.

3.2 Bayesian Models

The general framework just described cannot be applied directly to concrete problems and questions. To do so, the agent has to specify (1) the relevant variables and (2) their relations. This may involve a considerable amount of modeling, as it isn't always clear which variables are the relevant ones.[7] Besides, different methodological values may conflict. One may, for example, favor a rather

simple, intuitive, and understandable model. This can be achieved by taking into account only a small number of variables. Sometimes it is also possible to effectively combine various variables into one macro-variable. On the other hand, one might also want to account for the details of a reasoning scenario, and additional propositions might be needed for this. The modeler has to decide what to do. The same holds for the assumed relations between the variables. Are two variables really strictly (conditionally) independent? This may be controversial, and strict independencies are hard to come by.

However, once the modeler has decided what the relevant variables and their relations are, Bayes nets come in and help with the representation and the computations.[8] We will see that certain independence assumptions can do a lot of work and, if properly motivated, reduce the amount of subjectivity that one might complain about in a Bayesian model, due to the considerable freedom one otherwise has to fix a prior probability distribution. The structural constraints imposed in a Bayes net alone already contribute a great deal to the solution of a problem, and in the study of Bayesian rationality, probabilistic models often have Bayes nets as integral parts. Besides, once the Bayes net is fixed, the problem or question that prompted the construction of the model can be addressed by powerful mathematical machinery.

It has long been noted that models are an indispensable part of science.[9] They have important (pragmatic and epistemic) functions in the research process, and this also holds for the Bayes net models we are considering in this chapter. Here is a (nonexhaustive) list of the functions of Bayes net models in rationality research:

- Models help to *apply* the (Bayesian) theory. We mentioned already that not much follows from the general framework for a particular problem or question; it requires the model and the specification of details.
- Models help to *test* the (Bayesian) theory. Once the model is constructed, its consequences can be confronted with empirical data. Note, however, that Bayes net models are *normative models*, which raises additional problems (see Colyvan, 2013; Titelbaum, 2021).
- Models help to *solve* concrete problems, as many examples show (see also the illustrations in figure 4.2.4).
- Models *accommodate and explain* experimental data.
- Models help the researcher to *reconstruct and analyze* various reasoning and argumentation schemes and check their rationality.
- Bayes net models provide a *compact and intuitive representation* of a reasoning and argumentation situation and help with the bookkeeping of the variables and their relations.

3.3 Modeling Confirmation Scenarios

Large parts of the literature on BCT are only concerned with scenarios involving two variables—one representing the hypothesis (*H*) and the other representing the evidence (*E*). However, while much can be learned from focusing on simple scenarios (see, e.g., Earman, 1992; Howson & Urbach, 2006), it is clear that actual confirmation scenarios are much more complex and may raise new and intricate issues that do not show up in scenarios involving only two variables. If we want to reconstruct and analyze more complex confirmation scenarios, then Bayes nets prove to be extremely helpful. Without the use of Bayes nets, it is hard to keep track of the various dependencies and independencies and to properly evaluate what is going on, as the following example illustrates.

We consider the *testing of a hypothesis with partially reliable measurement instruments*. To begin with, we consider a situation where the evidence is uttered by a partially reliable information source. This information source could be a measurement instrument (which outputs a binary result, e.g., a positive or negative X-ray) or the testimony of a witness (e.g., in a murder case), again modeled as a binary variable, such as "I saw the suspect on the crime scene" or "I didn't see the suspect."

Such scenarios can be modeled by fixing the rates of false positives and false negatives of the information source. The rate of false positives (f_p) measures how often an instrument outputs "true" when in fact the hypothesis is false. Similarly, the rate of false negatives (f_n) measures how often an instrument outputs "false" when in fact the hypothesis is true. Ideally, both rates are zero, but in realistic cases, this is almost never the case. To proceed, it is easy to see that f_p and f_n are related to the likelihoods mentioned above: $f_p = P(\mathrm{E} \mid \neg \mathrm{H}) = q$ and $f_n = P(\neg \mathrm{E} \mid \mathrm{H}) = 1 - p$. While these likelihoods (or rates) are often available (e.g., in medical testing), this is not always the case (e.g., for witness reports). This suggests that one might want to construct more complex models for the information-gathering process of an agent.

Here is a simple example. Consider an agent who entertains the hypothesis H ("Paul is the murderer"). She then receives a witness report Rep (which is the evidence) to the effect that Paul indeed is the murderer. The agent then assumes—and this is a modeling assumption—that

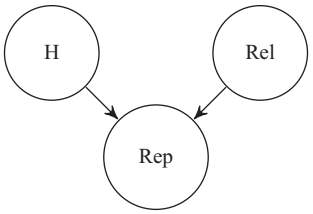


Figure 4.2.4
A Bayes net representing the test of a hypothesis with a partially reliable information source.

the witness is either reliable or not: if the witness is reliable, then she is fully reliable and therefore always tells the truth. That is, if Paul is the murderer, she reports that he is the murderer, and if Paul is not the murderer, she reports that Paul is not the murderer. However, if she is not reliable, then she is a *randomizer*, that is, she reports with a certain probability a (the so-called randomization parameter) that Paul is the murderer, independently of whether or not Paul is the murderer. Finally, the agent assumes that it is uncertain whether or not the witness is reliable and assigns a prior probability to the proposition Rel ("The witness is reliable").

Next we represent this situation with the ("collider") Bayes net in figure 4.2.4 and set $P(\mathrm{H}) = h$, $P(\mathrm{Rel}) = r$, and

$$P(\mathrm{Rep} \mid \mathrm{H}, \mathrm{Rel}) = 1, \qquad P(\mathrm{Rep} \mid \neg\mathrm{H}, \mathrm{Rel}) = 0,$$
$$P(\mathrm{Rep} \mid \mathrm{H}, \neg\mathrm{Rel}) = a, \qquad P(\mathrm{Rep} \mid \neg\mathrm{H}, \neg\mathrm{Rel}) = a.$$

With this, we calculate $P'(\mathrm{H}) = P(\mathrm{H} \mid \mathrm{Rep})$, that is, the posterior probability of the hypothesis after receiving a positive witness report:

$$\begin{aligned} P'(\mathrm{H}) &= \frac{P(\mathrm{H}, \mathrm{Rep})}{P(\mathrm{Rep})} = \frac{\sum_{Rel} P(\mathrm{H}, Rel, \mathrm{Rep})}{\sum_{H, Rel} P(H, Rel, \mathrm{Rep})} \\ &= \frac{\sum_{Rel} P(\mathrm{H})\, P(Rel)\, P(\mathrm{Rep} \mid \mathrm{H}, Rel)}{\sum_{H, Rel} P(H) P(Rel) P(\mathrm{Rep} \mid H, Rel)} \\ &= \frac{h(r + a\bar{r})}{h(r + a\bar{r}) + \bar{h}a\bar{r}} = \frac{h(r + a\bar{r})}{hr + a\bar{r}}. \end{aligned} \tag{2}$$

It is interesting to note that this model can be simulated by a two-variable model (using the variables H and Rep) with the corresponding likelihoods $p := P(\mathrm{Rep} \mid \mathrm{H}) = r + a\bar{r}$ and $q := P(\mathrm{Rep} \mid \neg\mathrm{H}) = a\bar{r}$. From this, one also sees that $p > q$ (for $r > 0$). Hence, as expected, Rep always confirms H. From equation 2, we obtain that $P'(\mathrm{H}) = P(\mathrm{H})$ for $r = 0$. Likewise, for $r = 1$, we find that $P'(\mathrm{H}) = 1$. This makes sense: if a perfectly reliable information source tells us that H is true, then H is true.

Because this three-variable model can be simulated by a two-variable model, it is not strictly necessary to study it in detail. However, the model allows us to reduce the likelihoods p and q to some parameters that are easier to grasp and interpret (namely, a and r). The model therefore provides (or suggests) a *mechanism* that generates the likelihoods, but once these likelihoods are known, one can proceed with the two-variable model.

Things get more interesting when one studies more complex scenarios. Let us assume, for instance, that the agent receives two positive reports from two partially reliable information sources. We can then distinguish two scenarios and ask which of them provides more confirmation for the hypothesis in question. In the first scenario, the two reports are independent from each other. This is modeled by fixing two root nodes Rel_1 and Rel_2 (see figure 4.2.5). In the second scenario, we assume that the reports are dependent and model this by working with only one root node Rel, which is a parent of both Rep_1 and Rep_2 (see figure 4.2.6). One would expect that the first scenario provides more confirmation, ceteris paribus, as the two reports are independent. Here, the ceteris paribus clause makes sure that the priors of H and Rep (and of Rep_1 and Rep_2, respectively) are the same and that the randomization parameter is the same in both scenarios. But is this really the case? Do independent positive reports always result in more confirmation? A detailed analysis shows that the answer to this question is "no": the second scenario provides more confirmation if the values of a and r are sufficiently small.[10]

Extending and generalizing these ideas, Landes (2020) provides a systematic assessment of the *variety-of-evidence thesis*. This is the claim that more varied evidence confirms more strongly than less varied evidence. Landes

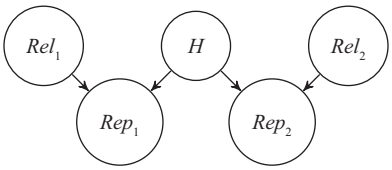


Figure 4.2.5
A Bayes net representing the test of a hypothesis with two independent instruments.

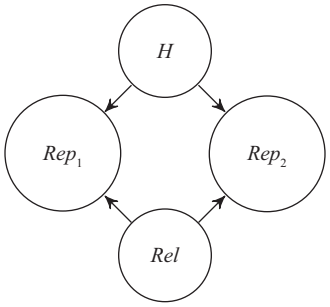


Figure 4.2.6
A Bayes net representing the test of a hypothesis with two dependent instruments.

explores in detail under which conditions the thesis holds. Further applications of the simple witness model described above can be found in the psychological literature (see, e.g., Hahn, Merdes, & von Sydow, 2018; Harris, Hahn, Madsen, & Hsu, 2016).

There are many other applications of Bayes nets in BCT. For instance, Dardashti, Hartmann, Thébault, and Winsberg (2019) investigate whether so-called analogue simulations can be used to confirm a theory. The (very rough) idea is this: A theory T predicts a phenomenon φ, but for certain (practical) reasons, the prediction cannot be directly tested. However, it turns out that there is another theory T′, which is in important respects analogous to T and predicts the analogous phenomenon φ'. Fortunately, the prediction of φ' can be tested. The observation of φ' then confirms T′, but does it also confirm T (as some authors argue for some cases)? Using the machinery of Bayes nets, Dardashti et al. (2019) show that this is indeed the case, given that certain conditions hold. Further applications concern the use of Bayes nets in legal reasoning. For example, Lagnado and collaborators (e.g., Connor Desai, Reimers, & Lagnado, 2016; Fenton, Neil, Yet, & Lagnado, in press) use Bayes nets to represent complex legal scenarios, demonstrating the power and the flexibility of the approach (see also chapter 11.4 by Prakken, this handbook).

4. Bayesian Reasoning and Argumentation

Confirmation theory and the analysis of confirmation scenarios are an important focus of Bayesianism. Bayesianism provides a clear criterion for when it is rational for an agent to claim that a piece of evidence supports a given hypothesis, and Bayes nets help to apply this theory to concrete cases so that the theory can be put to work. At this point, however, it is important to note that epistemic rationality comprises more than confirmation theory. In many cases of reasoning and argumentation, it is not immediately clear that such scenarios can be reconstructed (or represented) as confirmation scenarios. Interestingly, though, often it is possible. We illustrate this point with the sketch of a Bayesian account of argumentation (see also chapter 5.5 by Hahn & Collins and chapter 5.6 by Woods, both in this handbook).

The starting point of Bayesian argumentation is the observation that the premises and conclusions in typical real-life arguments are uncertain to the agent. We are more or less convinced of the premises of an argument, given the evidence we have for them (and our background knowledge), and this uncertainty transfers to the argument's conclusion. This holds whether the underlying argument scheme is valid or not. Hahn and Oaksford (2007) have observed that even so-called logical fallacies (such as the argument scheme Denying the Antecedent) can be powerful in the sense that they can increase an agent's degree of belief in the conclusion of the argument. Hahn and Oaksford argue convincingly that it is not only the logical *structure* that is important when it comes to assessing the strength of an argument but also the *content* of the premises (see also Hahn & Hornikx, 2016). Generalizing from these insights, Eva and Hartmann (2018a) develop a general theory of Bayesian argumentation according to which "Argumentation is learning." This slogan connects argumentation with confirmation, and we will sketch the corresponding theory now. Some of its applications and further developments are discussed in the following subsections.

We consider an agent (agent 1) who entertains a set of propositions with a prior probability distribution P defined over it. This can be represented by a Bayes net. Now another agent (agent 2) wants to convince agent 1 of some proposition. She decides to do so in an indirect way by manipulating the beliefs of agent 1 about the premises of an argument. More specifically, agent 2 aims at getting agent 1 to increase her degree of belief in some of the premises. To make sure that her overall degrees of belief are coherent, agent 1 then updates on that new information, which leads to a new probability of the conclusion. Hence the slogan "Argumentation is learning." If the new probability of the conclusion is greater than the old one, then the argument has some force; if not, then not. Note that the possible logical relationship between the premises and the conclusion plays a role in the updating process. This has been noted long ago by Suppes (1966) and Adams (1996).

As an illustration, consider modus ponens, that is, the rule

$$\frac{\begin{array}{c}A \rightarrow C \\ A\end{array}}{C}$$

and assume that the agent has a prior probability distribution P defined over the two variables. To reconstruct the argument in Bayesian terms, we assume that the agent learns the premises of the argument with certainty. We can then use Conditionalization to compute the new probability of C. Representing the conditional $A \rightarrow C$ by the corresponding material conditional $\neg A \vee C$ finally yields

$$P'(C) = P(C \mid A, \neg A \vee C) = \frac{P(A, C, \neg A \vee C)}{P(A, \neg A \vee C)} = 1.$$

Hence the argument succeeds. As a result, the agent increases the probability of the conclusion to 1 and thus becomes certain that C is true. In less ideal circumstances, for example, if the probability of the minor premise does not increase to 1 but to some value smaller than 1, one finds that the probability of the conclusion in a modus

ponens argument increases but never reaches 1. Interestingly, for valid arguments, this always holds: if the conditional is learnt with certainty and the probability of the minor premise increases, then the probability of the conclusion of the argument increases. This is not the case for invalid arguments, which is a reason to prefer valid ones (see Adams, 1996). If the argument scheme is invalid, then it depends on the prior probability distribution of the agent whether or not the probability of the conclusion increases. We will come back to this in subsection 4.3. But before, let us illustrate Bayesian argumentation with two examples that also illustrate the importance of Bayes nets.

4.1 The No-Alternatives Argument

Consider the following argument:

> A scientist entertains a theory H that satisfies several desirable conditions. Unfortunately, however, the theory cannot be tested empirically because it is mathematically too difficult to derive predictions from the theory or it is impossible to experimentally test the predictions of the theory. At this point, the scientist argues as follows: "Look, my colleagues and I tried hard to find an alternative to H that also satisfies the desirable conditions, but despite a lot of effort and brain power, we did not succeed. This supports the claim that H is true." That is, the scientist argues that an observation about the performance of the scientific community can provide a reason in favor of a scientific theory.

This is the No-Alternatives Argument (NAA). Quite recently, it has been put forward by defenders of string theory, which is a candidate for the most fundamental theory of physics that arguably provides a unified account of all four fundamental forces of nature (i.e., gravity, electromagnetism, and the weak and strong nuclear forces). Unfortunately, string theory lacks direct empirical confirmation, and no one expects this situation to change in the foreseeable future. This raises the question why scientists stick to a theory that is not (and perhaps never will be) confirmed by empirical data.

In his recent book *String Theory and the Scientific Method* (2013), R. Dawid suggests that the NAA is convincing, given certain conditions, and R. Dawid, Hartmann, and Sprenger (2015) provide a Bayesian analysis of the NAA. We present a slightly simplified version of their reconstruction and consider the following argument:

P_1: Theory H satisfies several desirable conditions.

P_2: Despite a lot of effort, the scientific community has not yet found an alternative to H that also satisfies these conditions.

C: Hence we have one good reason in favor of H.

Let F be the proposition "The scientific community has not yet found an alternative to H." One then has to show that F confirms H. Using BCT, we therefore have to show that $P'(\mathrm{H}) = P(\mathrm{H}\,|\,\mathrm{F}) > P(\mathrm{H})$. Note that F is neither a deductive nor an inductive consequence of H. Hence there cannot be a direct probabilistic dependence between the corresponding propositional variables. It is therefore natural to look for a third variable that facilitates the dependence. One possibility is a "common cause" structure with the two "effects" H and F and a so-far-unknown "common cause" variable Y that screens off H and F from each other. Such a structure seems appropriate as it renders H and F dependent if the agent does not know the value of Y, and H and F become independent once this value is known. But what could this variable be? We need to find another *active variable* about which the agent has beliefs and which is related to H and F in a simple and easy-to-justify way. (We certainly want to avoid that the result depends too much on the idiosyncrasies of the specific model.) R. Dawid et al. (2015) suggest the introduction of the multivalued variable Y, which has the following values:

Y_i: There are i distinct alternative theories that satisfy the desirable conditions.

Here, i runs from 0 to some maximal value N. Each Y_i is a statement about the number of *existing* theories that satisfy the same conditions as H. It is easy to see that Y screens off F from H: Once we know the value of Y, learning F does not tell us anything new about the probability of H. To assess it, all that matters is that we know how many equally suitable candidate theories there are. Y facilitates the probabilistic dependence between F and H, since if there *is* only a small number of alternative theories, this would provide an explanation for why the scientists have not yet found any (i.e., it would explain F). In addition, if there are only a few alternative theories, this should also probabilistically affect our belief in the available theory. We finally arrive at the Bayes net depicted in figure 4.2.7 and formulate the following four plausible conditions (see unnumbered list below):

NAA1: $\forall i \geq 0$, $y_i := P(\mathrm{Y}_i) \in [0, 1)$.

NAA2: $h_i := P(\mathrm{H}\,|\,\mathrm{Y}_i)$ are monotonically decreasing in i.

NAA3: $f_i := P(\mathrm{F}\,|\,\mathrm{Y}_i)$ are monotonically decreasing in i.

NAA4: H and F are conditionally independent given Y, that is, $P(\mathrm{H}\,|\,\mathrm{F}, \mathrm{Y}_i) = P(\mathrm{H}\,|\,\mathrm{Y}_i)$.

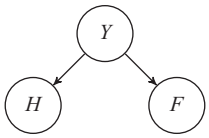


Figure 4.2.7
A Bayes net depicting the No-Alternatives Argument.

It can then be shown that the following theorem holds:

> *Theorem 1.* Let P be a probability distribution satisfying the conditions NAA1–NAA4. Then F confirms H, that is, $P(\mathrm{H} \mid \mathrm{F}) > P(\mathrm{H})$, iff there exists a pair (i, j) with $i > j$ such that (1) $y_i\, y_j > 0$, (2) $f_i < f_j$, and (3) $h_i < h_j$.

This theorem allows for an assessment of the NAA. One may, for example, question NAA1. Are we really uncertain about the number of alternatives? Doesn't the *underdetermination thesis* in philosophy of science tell us that the number of alternatives to a given theory is always infinite? If this is so, then a rational agent should set $y_\infty = 1$ and the NAA does not go through. A defender of the NAA will therefore have to argue why y_∞ (and all other y_i, for that matter) should be assigned values smaller than 1. It has to be shown, then, that an agent can never be certain about the number of appropriate alternatives to a given theory.

Here is another issue. The NAA relies on how well F probes the variable Y. Note that there can be several complicating factors. For one, there might be an alternative explanation for why the scientific community has not yet found an alternative. For instance, R. Dawid et al. (2015) introduce an additional node D representing the difficulty of finding an alternative theory. The observation of F then may only confirm D (i.e., that it is very difficult to find an alternative theory). However, D is probabilistically independent of H, and hence the observation of F may not confirm the theory in question.

Note also that the NAA relies on it being possible to establish F in the first place. However, it is a nontrivial task to find agreement among the members of the scientific community about the existence or nonexistence of alternative theories. And even if there is agreement, this is only probative of the number of existing alternatives (i.e., the value of the variable Y), provided that the scientific community has attempted to explore the space of alternative theories and has considered all problems one may encounter in this endeavor.

We conclude that the formal reconstruction of the NAA using Bayes nets helps the reasoner to put all assumptions on the table and highlights issues that need to be addressed to make the argument, if at all possible, convincing.

4.2 No Reason For Is a Reason Against

Next, consider the following argument:

> You are interested in the question whether or not God exists. To address this question, you consider all arguments for the existence of God you can find in the literature. After a careful examination of them, you come to the conclusion that none of them is convincing. From *this* observation, you conclude that you have a reason *against* the existence of God because *no reason for is a reason against*.

This is the No Reason For Argument (NRF). While this reasoning may have some plausibility, it is not clear whether the argument is a good argument. To find out, we follow Eva and Hartmann (2018b) and provide a Bayesian reconstruction of the NRF, which proceeds analogously to the NAA and allows us to ask whether learning the premises of the argument increases the probability of its conclusion.

We consider the hypothesis H and introduce the proposition F, which says "I have not yet found a good argument in favor of H." Furthermore, let Y be a propositional variable whose values are the propositions Y_i: "There are exactly i good arguments in favor of H" ($i \geq 0$). It is plausible that an agent can be uncertain about the value of Y. Many rational agents would surely plead ignorance as to whether or not there are any undiscovered good arguments for the existence of God (and, if so, how many). Again, knowledge of the value of Y renders F independent of H: if I know that there are five good arguments for the existence of God, then the fact that I haven't yet found any one of them should be irrelevant to my belief in the existence of God. Furthermore, if I learn that there are more good arguments for God's existence than I previously thought, then that should raise my degree of belief in the existence of God. Finally, the more arguments there are for God's existence, the more likely it is that I find one. These considerations motivate the following conditions (see unnumbered list below):

NRF1: $\forall i \geq 0,\ y_i := P(\mathrm{Y}_i) \in [0, 1)$.

NRF2: $h_i := P(\mathrm{H} \mid \mathrm{Y}_i)$ are monotonically increasing in i.

NRF3: $f_i := P(\mathrm{F} \mid \mathrm{Y}_i)$ are monotonically decreasing in i.

NRF4: H and F are conditionally independent on Y, that is, $P(\mathrm{H} \mid \mathrm{Y}_i, \mathrm{F}) = P(\mathrm{H} \mid \mathrm{Y}_i)$.

NRF1–NRF4 are structurally near-identical to the conditions imposed on the NAA. The corresponding Bayes net model is also analogous. The only difference is that the h_i are now monotonically increasing in i. Accordingly, the proof for the following theorem is analogous to the proof of theorem 1.

Theorem 2. Let P be a probability distribution satisfying the conditions NRF1–NRF4. Then F disconfirms H, that is, $P(\mathrm{H}|\mathrm{F}) < P(\mathrm{H})$, iff there exists a pair (i, j) with $i > j$ such that (1) $\gamma_i\, \gamma_j > 0$, (2) $f_i > f_j$, and (3) $h_i < h_j$.

We contend that theorem 2 constitutes, under certain special circumstances, a full Bayesian vindication of the NRF argument. On the basis of this reconstruction, one can analyze how good specific NRF arguments are (for details, see Eva & Hartmann, 2018b).

4.3 Toward a General Theory of Reasoning and Argumentation

The examples given in the last two subsections illustrate the slogan "Argumentation is learning" mentioned above. Both examples involved the learning of a proposition, and the learning was modeled using Conditionalization. Note, however, that many argument schemes involve indicative conditionals of the form "If A, then C" as premises, which are notoriously difficult to deal with. In the introduction to this section, we represented an indicative conditional by the corresponding material conditional $\neg \mathrm{A} \vee \mathrm{C}$. This was convenient, as it allowed us to condition on it. However, the material conditional faces many problems, and it is even debated whether indicative conditionals can be modeled as propositions at all (for details, see Douven, 2018, and chapter 6.1 by Starr, this handbook). It is therefore advisable to look for a more general updating method that can be applied to both propositional and nonpropositional evidence. A general theory of Bayesian argumentation cannot work without such a method. To address this problem, Eva et al. (2020) develop the *distance-based approach to Bayesianism* that builds on the abovementioned principle of Conservativity. Learning the indicative conditional "If A, then C" from a perfectly reliable information source then suggests the constraint $P'(\mathrm{C}\,|\,\mathrm{A}) = 1$. The full posterior probability distribution P' follows by minimizing a suitable distance measure (such as the Kullback–Leibler divergence) between P' and P. Eva and Hartmann (2018a) apply this proposal to Bayesian argumentation and show how Bayes nets can be used to model, for example, disablers in an argument.

5. Outlook

In this chapter, we have introduced the theory of Bayes nets, shown how Bayes net models can be constructed within the Bayesian framework, and presented a number of examples from the philosophy and psychology of reasoning and argumentation that illustrate the power of the approach. However, Bayes nets have also been used in other areas of rationality research (such as the study of decision making), and there is no doubt that a whole range of further problems can be addressed with this methodology.

It is also worth noting that Bayes nets can be used with other theories of uncertainty as well. For example, Spohn (2012, chapter 7) shows that ranking theory satisfies the semi-graphoid axioms. Other authors investigate conditional independence structures for imprecise probabilities (see, e.g., Cozman, 2012; Halpern, 2003, chapter 4). Hence, nothing hinges on the Bayesian framework that we used in this chapter. At the same time, it is true that most applications of Bayes nets in rationality research studied so far presuppose the Bayesian framework, which is easy to use and has a solid normative foundation. It would be interesting to construct analogous models in other frameworks and to compare the resulting analyses with the corresponding Bayesian analyses (see also Colombo, Elkin, & Hartmann, in press). This will lead to further progress in rationality research.

Acknowledgments

Thanks to my collaborators Richard Dawid, Benjamin Eva, and Jan Sprenger and to the editors of this volume for their patience and valuable feedback.

Notes

1. For a textbook exposition, see Darwiche (2014, section 4.4) and Pearl (1988, section 3.1).

2. We follow the convention, adopted, for example, in Bovens and Hartmann (2003), to represent propositional variables in italics and their values in roman script. For instance, the variable A has the values A and $\neg$A. Here and in the remainder, we also use the shorthand notation $P(\mathrm{A}, \mathrm{B})$ for $P(\mathrm{A} \wedge \mathrm{B})$.

3. If A_j is a root node, then $\mathrm{pa}(\mathrm{A}_j)$ is the empty set ($\varnothing$) and $P(\mathrm{A}_j\,|\,\mathrm{pa}(\mathrm{A}_j)) = P(\mathrm{A}_j)$.

4. We sometimes use causal language when talking about various Bayes nets. And indeed, causal intuitions help when it comes to construct a Bayes net model, which typically respects

the "causal direction." For instance, we always draw an arrow from the hypothesis variable to the corresponding evidence variable. Note, however, that all we need for the applications discussed in this chapter is that a Bayes net represents a joint probability distribution. Hence, the causal language is, strictly speaking, only of heuristic value here (see chapter 7.1 by Pearl, this handbook).

5. Section 5 of this volume surveys the various contenders. See also chapter 4.7 by Dubois and Prade (this handbook) and Halpern (2003). Oaksford and Chater (2007) discuss topics in the psychology of reasoning from a Bayesian point of view. Hájek and Hartmann (2010) and Hartmann and Sprenger (2010) survey Bayesian epistemology, and Sprenger and Hartmann (2019) investigate various topics in the philosophy of science from a Bayesian point of view.

6. "Relevant" means relevant for the specific problem or question the agent is interested in.

7. It would be very helpful to have an automated way to extract the relevant variables and their relations from potentially large data sets. For attempts in this direction, see Chalupka, Bischoff, Perona, and Eberhardt (2016).

8. The choice of the variables, and which relations between them one assumes, may also be suggested by considerations about which Bayes net provides a good representation.

9. See Frigg and Hartmann (2016/2020) for a survey of the corresponding philosophy of science literature.

10. For details and an explanation, see Bovens and Hartmann (2003).

References

Adams, E. W. (1996). *A primer of probability logic* (Synthese Library, Vol. 86). Boston, MA: Reidel.

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford, England: Clarendon Press.

Chalupka, K., Bischoff, T., Perona, P., & Eberhardt, F. (2016). Unsupervised discovery of El Niño using causal feature learning on microlevel climate data. In *Proceedings of the 32nd Conference on Uncertainty in Artificial Intelligence* (pp. 72–81). Arlington, VA: AUAI Press.

Colombo, M., Elkin, L., & Hartmann, S. (in press). Being realist about Bayes, and the predictive processing theory of mind. *British Journal for the Philosophy of Science*.

Colyvan, M. (2013). Idealisations in normative models. *Synthese*, *190*, 1337–1350.

Connor Desai, S., Reimers, S., & Lagnado, D. (2016). Consistency and credibility in legal reasoning: A Bayesian network approach. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 626–631). Austin, TX: Cognitive Science Society.

Cozman, F. G. (2012). Sets of probability distributions, independence, and convexity. *Synthese*, *186*(2), 577–600.

Crupi, V. (2016). Confirmation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from https://plato.stanford.edu/archives/spr2020/entries/confirmation/

Dardashti, R., Hartmann, S., Thébault, K., & Winsberg, E. (2019). Hawking radiation and analogue experiments: A Bayesian analysis. *Studies in History and Philosophy of Modern Physics*, *67*, 1–11.

Darwiche, A. (2014). *Modeling and reasoning with Bayesian networks*. Cambridge, England: Cambridge University Press.

Dawid, A. P. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society, Series B*, *41*, 1–31.

Dawid, R. (2013). *String theory and the scientific method*. Cambridge, England: Cambridge University Press.

Dawid, R., Hartmann, S., & Sprenger, J. (2015). The no alternatives argument. *British Journal for the Philosophy of Science*, *66*(1), 213–234.

Diaconis, P., & Zabell, S. L. (1982). Updating subjective probability. *Journal of the American Statistical Association*, *77*(380), 822–830.

Douven, I. (2018). *The epistemology of indicative conditionals: Formal and empirical approaches*. Cambridge, England: Cambridge University Press.

Earman, J. (1992). *Bayes or bust? A critical examination of Bayesian confirmation theory*. Cambridge, MA: MIT Press.

Eva, B., & Hartmann, S. (2018a). Bayesian argumentation and the value of logical validity. *Psychological Review*, *125*(5), 806–821.

Eva, B., & Hartmann, S. (2018b). When no reason for is a reason against. *Analysis*, *178*(3), 426–431.

Eva, B., Hartmann, S., & Rafiee Rad, S. (2020). Learning from conditionals. *Mind*, *129*(514), 461–508.

Fenton, N., Neil, M., Yet, B., & Lagnado, D. (2020). Analyzing the Simonshaven case using Bayesian networks. *Topics in Cognitive Science*, *12*(4), 1092–1114.

Frigg, R., & Hartmann, S. (2020). Models in science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from https://plato.stanford.edu/archives/spr2020/entries/models-science/

Giere, R. N. (1990). *Explaining science: A cognitive approach*. Chicago, IL: University of Chicago Press.

Hahn, U., & Hornikx, J. (2016). A normative framework for argument quality: Argumentation schemes with a Bayesian foundation. *Synthese*, *193*(6), 1833–1873.

Hahn, U., Merdes, C., & von Sydow, M. (2018). How good is your evidence and how would you know? *Topics in Cognitive Science*, *10*(4), 660–678.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, *114*(3), 704–732.

Hájek, A., & Hartmann, S. (2010). Bayesian epistemology. In J. Dancy, E. Sosa, & M. Steup (Eds.), *A companion to epistemology* (2nd ed., pp. 93–106). Oxford, England: Wiley-Blackwell.

Halpern, J. Y. (2003). *Reasoning about uncertainty*. Cambridge, MA: MIT Press.

Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2016). The appeal to expert opinion: Quantitative support for a Bayesian network approach. *Cognitive Science*, *40*(6), 1496–1533.

Hartmann, S., & Sprenger, J. (2010). Bayesian epistemology. In S. Bernecker & D. Pritchard (Eds.), *The Routledge companion to epistemology* (pp. 609–620). London, England: Routledge.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach*. London, England: Open Court.

Huber, F. (2007). Confirmation and induction. In *Internet encyclopedia of philosophy*. Retrieved from https://iep.utm.edu/conf-ind/

Landes, J. (2020). Variety of evidence. *Erkenntnis*, *85*, 183–223.

Oaksford, M., & N. Chater (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, England: Oxford University Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Francisco, CA: Morgan-Kaufmann.

Spohn, W. (1980). Stochastic independence, causal independence, and shieldability. *Journal of Philosophical Logic*, *9*, 73–99.

Spohn, W. (2012). *The laws of belief: Ranking theory and its philosophical applications*. Oxford, England: Oxford University Press.

Sprenger, J. (2011). Hypothetico-deductive confirmation. *Philosophy Compass*, *6*(7), 497–508.

Sprenger, J., & Hartmann, S. (2019). *Bayesian philosophy of science*. Oxford, England: Oxford University Press.

Suppes, P. (1966). Probabilistic inference and the concept of total evidence. In J. Hintikka & P. Suppes (Eds.), *Aspects of inductive logic* (pp. 49–65). Amsterdam, Netherlands: North Holland.

Titelbaum, M. G. (2021). Normative modeling. Retrieved from http://philsci-archive.pitt.edu/18670

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Knauff, Markus, editor. | Spohn, Wolfgang, editor.
Title: The handbook of rationality / edited by Markus Knauff and Wolfgang Spohn.
Description: Cambridge : The MIT Press, 2021. | Includes bibliographical references and index.
Identifiers: LCCN 2020048455 | ISBN 9780262045070 (hardcover)
Subjects: LCSH: Reasoning (Psychology) | Reason. | Cognitive psychology. | Logic. | Philosophy of mind.
Classification: LCC BF442 .H36 2021 | DDC 153.4/3—dc23
LC record available at https://lccn.loc.gov/2020048455