4.7 Representing Belief: Beyond Probability and Logic

Didier Dubois and Henri Prade

Summary

This chapter surveys recent approaches to the representation of belief. There is a clash between the notion of degree of belief in the subjective probability tradition and the idea of certainty as accepted belief, often couched in the language of logic, especially modal logic. The attempt to consider degrees of certainty finds its origin in the works of Francis Bacon, often opposed to the one of Pascal. However, by reviewing some more recent trends in the representation of uncertainty, such as possibility theory, ranking theory, evidence theory, and imprecise probability, one may argue that these novel approaches try to bridge the gap between the two traditions, even if dropping some favorite properties on the way, such as the additivity of degrees of belief and the adjunction law for accepted beliefs.

1. A Clash of Traditions

There are two traditions for representing belief, one based on probability, one based on modal logic. The probability tradition goes back to Pascal and other scientists of his time, who tried to measure our confidence in events (today we also speak of "propositions" in the logical sense). In those times, a distinction was made between so-called chances and probabilities (Shafer, 1978). The word "chance" was dedicated to repeatable events and measured via a combinatorial count of the frequencies of occurrence, assuming elementary events had equal chances to occur. The name "probability" was understood as a subjective notion, typically the confidence to be granted to a testimony. The latter view is still prevalent in important texts from the 18th century, such as d'Alembert and Diderot's encyclopedia.[1] The concept of subjective probability was to be taken up again in the 20th century by scholars like Ramsey (1926/1931) and de Finetti (1937). They tried to justify why degrees of belief should be additive. However, additivity can only be obtained by introducing severe restrictions on the model of graded belief namely, it essentially prevents capturing the notion of ignorance. The reader is referred to the rather recent compilation by Huber and Schmidt-Petri (2009) for a survey of approaches to degrees of belief.

In contrast, the modal logic tradition was developed in the 20th century, even if modal logic per se goes back to medieval philosophy and even to ancient Greek philosophy. Hintikka (1962) proposed a so-called epistemic logic that tried to define a rational logical framework for the idea of knowledge and its articulation with an all-or-nothing notion of belief. Namely, knowledge is essentially true belief, while a simply believed proposition is not supposed to be actually true. In this setting, modeling ignorance is obvious as you just express that you believe neither a proposition nor its negation. Besides, one essential feature of this approach is that when you know that each of two propositions is true, you also know that their conjunction is true, which is sometimes called the *adjunction axiom* (Kyburg, 1997). This axiom is also supposed to hold for beliefs, with the understanding that modal beliefs represent *accepted beliefs*, that is, propositions taken for granted to the point of reasoning with them as if they were true.

There is a clash of intuitions between the two views, because for them to match, one needs to explain the connection between graded belief and belief *simpliciter*. One way is to add a threshold to graded beliefs and to retain propositions above this threshold, but it is then almost impossible to preserve the adjunction axiom if we model graded beliefs by probabilities. This is the point made by the lottery paradox (Kyburg, 1961), discussed later in this chapter.

The aim of this chapter is to provide an overview of belief representations that relate, and go beyond, subjective probability and epistemic logic, in the sense that some of them are graded versions of the epistemic logic approach and all of them give up the additivity assumption of probability theory. Most of these approaches have been developed only since the middle of the 20th century, with a sudden acceleration from 1975 on. This chapter should be of particular interest to psychologists, because they often stick to the alternative "logic vs. probability" and are not always aware of the fact that there is a large area beyond.

2. Rationality Axioms for Set Functions Representing Belief

In this section, we survey mathematical models of belief and axioms that have been proposed to represent them. We consider the most general approaches and show that accepting the rule of adjunction considerably narrows the expressive power of representations of graded belief.

2.1 Degrees of Belief

We assume throughout the chapter that belief qualifies propositions, represented by subsets A, B, C, . . . of a set S of possible worlds or states of affairs. More precisely, A stands for the proposition that event A occurs or equivalently that some entity of interest, say, the result x of a measurement, lies in A. We do not distinguish between logically equivalent propositions. Quantifying (the possible lack of) belief is one instance of expressing uncertainty for an agent.

The most usual representation of uncertainty consists in assigning to each proposition or event A a number $g(A)$ in the unit interval [0, 1]. It evaluates the confidence (we deliberately use a more general term than "belief" or "plausibility") of the agent in the truth of the proposition "$x \in A$." This proposition can only be true or false by convention, even if the agent may ignore its truth-value. The following requirements sound natural:

$$g(\emptyset) = 0, \qquad g(S) = 1,$$

as you cannot have confidence in contradictions, and you should believe in tautologies. Monotonicity with respect to inclusion is also natural:

$$\text{If } A \subseteq B \text{ then } g(A) \leq g(B).$$

Indeed, if A is more specific than B in the wide sense (in other words, A implies B), a rational agent should not be more confident in A than in B: all things being equal, the more imprecise a proposition, the more certain it is. Under these properties, the function g is sometimes called a *capacity* (after Choquet, 1953), sometimes a *fuzzy measure* (after Sugeno, 1977). In order to stick to the uncertainty framework, it was also called a *confidence measure* (Dubois & Prade, 1988). Such a set function represents the epistemic state of an agent, that is, what an agent knows or thinks he or she knows (irrespective of whether it is true or not).[2]

An important special case of a confidence measure is the probability measure $g = P$, which satisfies the *additivity* property:

$$\text{If } A \cap B = \emptyset, \text{ then } P(A \cup B) = P(A) + P(B).$$

Probability measures have very often been used to represent degrees of belief, hence with a subjective flavor, as opposed to limits of relative frequencies of repeatable events, which account for random phenomena. Contrary to frequencies, belief may be attached to unique events. However, the additivity axiom sounds much more natural when modeling frequencies than for degrees of belief. Actually, the latter were not supposed to be additive in pioneering works by Bernoulli and Lambert (Shafer, 1978).

To recover additivity, scientists of the mid-20th century such as de Finetti (1937) interpreted $P(A)$ as the amount of money the agent is ready to pay for buying a lottery ticket that earns \$1 if the proposition A is true. Moreover, it is assumed that the price is fair, that is, the agent would accept to sell this ticket under the same conditions for the same price, and wants to avoid sure loss. This interpretation of degrees of belief entails the additivity axiom. Hence, to represent degrees of belief, probability measures should be used. Such probabilities are called "subjective" (see chapter 4.1 by Hájek & Staffel, this handbook). Bayesian probabilists consider nonadditive degrees of belief irrational, since they incur sure money loss.

Remark: Supplying precise degrees of beliefs sounds cognitively very demanding. It seems more natural for an agent expressing his or her knowledge to only represent the relative strength of confidence in various propositions, rather than trying to force him or her to deliver numerical evaluations. It is indeed easier to assert that one proposition is more credible than another than to assess a particular degree of belief (whose meaning is not always simple to grasp) or even to guess a frequency for each of them. The idea of representing uncertainty by means of partial order relations over a set of events dates back to Ramsey (1926/1931), de Finetti (1937), and Koopman (1940). They tried to find an ordinal counterpart to subjective probabilities. Later, philosophers of logic such as David Lewis (1973) have considered other types of relations, including comparative possibilities in the framework of modal logic. We omit discussing this literature for the sake of brevity.

2.2 Accepted Belief

Alternatively to measuring degrees of belief, one may want to consider propositions as beliefs accepted by an agent, if the latter is ready to reason as if such believed propositions were true. In particular, if A and B are beliefs, the conjunction $A \cap B$ is also a belief (a debated issue). But the empty set should not be a belief. Moreover, if A is a belief, and $A \subseteq B$, then B is a belief. In other words, a set of beliefs should be consistent and deductively closed.

Consider a Boolean-valued set function N from 2^S to $\{0, 1\}$ (here, 2^S is the power set of S, i.e., the set of all subsets of S) such that $N(A) = 1$ if A is believed and $= 0$ otherwise. It is clear that N is a confidence measure (monotonic under inclusion) that satisfies the law of "minitivity":

$$N(A \cap B) = \min(N(A), N(B)).$$

This function is a special case of necessity measures (Dubois & Prade, 1988). It models the idea of certainty. It is easy to check that in the finite case, there exists a nonempty subset $E = \bigcap\{A\colon N(A) = 1\}$, and that $N(A) = 1$ if and only if $E \subseteq A$. The set E contains all states the agent does not consider impossible, given his or her beliefs, which is the simplest representation of an agent's epistemic state.[3] In other words, the function N describes propositions that can be proved from the epistemic state E.

2.3 Extracting Beliefs from Confidence Measures

In order to relate accepted beliefs to graded beliefs, one should extract accepted beliefs from a confidence function g. A natural way of proceeding is to define a belief as a proposition in which an agent has enough confidence. So we should define a positive belief threshold β such that A is a belief if and only if $g(A) \geq \beta > 0$. However, the closure of accepted beliefs under conjunction entails the following property:

Accepted belief postulate: *If* $g(A) \geq \beta$ *and* $g(B) \geq \beta$, *then* $g(A \cap B) \geq \beta$.

This requirement is very strong. As the set of accepted beliefs should not include the empty set, it is clear that one should have $\min(g(A), g(B)) < \beta$ if $A \cap B = \varnothing$. Worse, if the postulate holds for any positive threshold β, then it is clear that $g(A \cap B) \geq \min(g(A), g(B))$, but as g is monotonic under inclusion, it enforces the equality, which comes down to the statement that a capacity represents accepted beliefs if and only if $g(A \cap B) = \min(g(A), g(B))$, that is, g is a graded necessity measure, still denoted by N. Letting ι, a mapping from S to [0, 1], be the function defined by $\iota(s) = N(S \setminus \{s\})$ (the degree of belief that the actual state of affairs is not s), it is clear that $N(A) = \min_{s \notin A} \iota(s)$. The value $1 - \iota(s)$ can be interpreted as the degree of plausibility $\pi(s)$ of state s, where π is the membership function of a fuzzy epistemic state, usually called a "possibility distribution" (Zadeh, 1978). The set function Π from 2^S to $[0, 1]$ with

$$\Pi(A) = 1 - N(A^c) = \max_{s \in A} \pi(s),$$

where A^c is the set-theoretical complement of A in S, represents the degree of plausibility of A, measuring to what extent A is not ruled out by the agent. This setting is the one of *possibility theory* (Dubois & Prade, 1988). This means that possibility theory accounts for the notion of accepted belief.

Possibility theory was proposed by Lotfi A. Zadeh (1978) in the late 1970s to represent uncertain pieces of information expressed by fuzzy linguistic statements and was later developed in an artificial intelligence perspective (Dubois & Prade, 1988, 1998). Formally speaking, the proposal is quite similar to the one made almost 30 years before by the economist George L. S. Shackle (1949), who advocated and developed a nonprobabilistic view of uncertainty based on the idea of degree of potential surprise. The degree of potential surprise attached to proposition A can be modeled as $N(A^c)$, namely, the more you believe A^c, the more surprising you find the occurrence of A.

A necessity-like function was explicitly used by L. Jonathan Cohen (1977) under the name "Baconian probability." It can be traced back to the English philosopher Francis Bacon, while probability was investigated by Pascal and followers, as pointed out by Cohen (1980). It is devoted to the idea of provability in contrast to probability, and it perfectly fits necessity measures. Especially if you can prove A with some confidence, you cannot at the same time claim you can prove its negation, which makes the condition $\min(g(A), g(A^c)) = 0$ natural. This condition is satisfied by necessity measures. So, the condition $N(A) > 0$ expresses that A is an accepted belief, the absolute value of $N(A)$ expressing the strength of acceptance. Such Baconian probabilities, viewed as grades of certainty, are claimed to be more natural than probabilities, for instance, for use in legal matters. Deciding whether someone is guilty cannot be done using statistics, nor can it be based on betting probabilities: you must prove guilt using convincing dedicated arguments.

About a decade later, in the late 1980s, Wolfgang Spohn (1988) introduced the notion of *ordinal conditional functions*, now called *ranking functions*, as a basis for a dynamic theory of epistemic states. Ranking functions κ are a variant of potential surprise, taking values in the nonnegative integers, that is, $N_\kappa(A^c) = 2^{-\kappa(A)}$ is a degree of potential surprise. The theory of ranking functions (see chapter 5.3 by Kern-Isberner, Skovgaard-Olsen, & Spohn, this handbook) can be developed in parallel with possibility theory,[4] even though they were independently devised.

Subjective probability has been justified by Leonard Savage (1954) from first principles in the setting of decisions under uncertainty. This result has been very

This is a portion of the eBook at doi:10.7551/mitpress/11252.001.0001

important for popularizing probability as the natural way of representing degrees of belief. As it turns out, a similar approach has been carried out for possibility theory in the same act-based setting as the one of Savage, albeit assuming a finite state space S (for a detailed account, see Dubois, Prade, & Sabbadin, 2001). The approach leads to qualitative counterparts of expected utility and extends Wald's pessimistic and optimistic criteria to possibility distributions. They may be considered not discriminating enough and can be refined using special forms of expected utility that encode lexicographic refinements of min and max (leximin and leximax; Fargier & Sabbadin, 2005). For a survey, see Dubois, Fargier, Prade, and Sabbadin (2009); see also chapter 8.4 by Hill, this handbook.

2.4 Probability and Accepted Beliefs

It is clear that set functions representing accepted beliefs are at odds with probability measures supposed to capture rational degrees of belief. If we apply the threshold method to a probability function in order to recover accepted beliefs, then we fail to satisfy conjunctive closure since, however large the threshold β is,

$$P(A) > \beta \text{ and } P(B) > \beta \text{ do not imply } P(A \cap B) > \beta \text{ for all } A, B.$$

This point has been especially highlighted by Henry Kyburg (1961, 1997). He put forward the lottery paradox: if S contains the set of possible equiprobable outcomes of a chance game like buying a lottery ticket where only one wins, then, by making S big enough, we can make the probability of losing when betting on the chance outcome s, namely, $(|S| - 1)/|S|$, arbitrarily high. But if you buy all tickets, you are sure to win. This example questions the cogency of accepted belief understood as high-enough degree of belief. As a consequence, Kyburg proposed to give up the idea that accepted beliefs are deductively closed and constructed an appropriate logic accounting for this standpoint (Kyburg & Teng, 2012).

Another way out of the lottery paradox is suggested in Dubois, Fargier, and Prade (2004), namely, to restrict the set of probabilities. We look for probability measures that respect the adjunction rule, namely, for any context C and all A, B,

$$P(A|C) > P(A^c|C) \text{ and } P(B|C) > P(B^c|C) \text{ imply } P(A \cap B|C) > P(A^c \cup B^c|C).$$

This is the adjunction rule for $\beta = 0.5$. Such probability functions do exist, and they are called *big-stepped* probabilities (Benferhat, Dubois, & Prade, 1999) or also *atomic bound* probabilities (Snow, 1999); see also Leitgeb's (2014) stability theory. They have probability masses of the form $p_1 > p_2 > \cdots > p_n$, where $p_i = P(\{s_i\})$, such that $p_i > p_{i+1} + p_{i+2} + \cdots + p_n$ for all $i < n$, that is, they are the discrete counterpart of exponential probability functions. The same kind of result can be obtained for any acceptance threshold. It suggests that the notion of accepted belief is more consistent with probability theory for distributions that are strongly biased toward some specific outcome (the opposite of the uniform distributions used in the lottery paradox), that is, those for which commonsense beliefs can be entertained.

3. Qualitative Possibility Theory: Reasoning with Defeasible Accepted Beliefs

As mentioned before, there are two different traditions for doxastic reasoning: modal logic, which captures accepted beliefs by modalities, and set functions, which use probability. The fact that there is a class of set functions that also captures accepted beliefs suggests a bridge between the two traditions (Dubois et al., 2004). Let us also mention a third tradition, which models partial ignorance using a truth-functional many-valued logic (e.g., Kleene logic, which expresses partial knowledge about atomic propositions only). This severely limits its representation power (Ciucci & Dubois, 2013). We claim that qualitative possibility and necessity functions, valued in a bounded chain, offer a unified framework for these traditions. In particular, in this approach, accepted beliefs are defeasible, and it has close connections with nonmonotonic reasoning, counterfactual reasoning, and belief revision (Dubois et al., 2004; see also chapter 5.2 by Rott and chapter 6.1 by Starr, both in this handbook).

3.1 Boolean Set Functions and Modalities: A Simple Epistemic Logic

Consider a propositional language $\mathscr{L}$ where well-formed formulas $p, q, \ldots$ encode propositions. In artificial intelligence, a consistent set of formulas in propositional logic is often called a *belief base* and, if deductively closed, a *belief set*. In the syntax of propositional logic, you can express the fact that a proposition is believed, but there is no way to express that it is *not* believed. All you can express is that its negation is believed. Introducing a belief modality $\Box$ prefixing p improves the situation as we can now distinguish between $\neg\Box p$ (p is not believed) and $\Box\neg p$ (its negation is believed). The standard approach to modeling accepted beliefs is indeed modal logic, after Hintikka (1962). Note that the meaning of the modality $\Box$ is described by the axioms ruling the logic. It can range from a very loose interpretation ("the agent has received information that p is true") to a very strong one ("the agent knows that p, in the sense of true belief that p").

The usual axioms of doxastic logic are those of the modal logic KD45 (see chapter 5.1 by van Ditmarsch, this handbook), which presupposes a modal language $\mathscr{M}$ that extends $\mathscr{L}$ and allows for nested modalities (the symbol $\Rightarrow$ represents material implication):

(PL) All axioms of propositional logics for M-formulas.
(K) $\Box(p \Rightarrow q) \Rightarrow (\Box p \Rightarrow \Box q)$
(D) $\Box p \Rightarrow \neg\Box\neg p$
(S4) $\Box p \Rightarrow \Box\Box p$
(S5) $\neg\Box p \Rightarrow \Box\neg\Box p$

The inference rules are modus ponens and necessitation (if p is a theorem, deduce $\Box p$). The semantics is in terms of accessibility relations $R \subseteq S \times S$. The satisfaction of $\Box p$ at a state s based on relation R is defined by $sR \subseteq [p]$, where $sR = \{s' : (s, s') \in R\}$ and $[p]$ denotes the set of states where p is true.

This approach can be simplified in order to relate doxastic logic with the representation of accepted belief from the previous section:

- KD45 uses a complex language, which we can restrict by putting the modality $\Box$ only before propositions in $\mathscr{L}$.
- Axioms S4 and S5, which are often called "positive" and "negative introspection," seem to be there basically to make complex formulas of the language $\mathscr{M}$ equivalent to simpler ones without nested modalities (even if they have a philosophical meaning).

It may sound counterintuitive to evaluate doxastic formulas on a real state of affairs; it is more natural to do it on epistemic or doxastic states (nonempty subsets of S).

The propositional language $\mathscr{L}_\Box$, whose atomic variables are of the form $\Box p$ ($p \in \mathscr{L}$), is the simplest language for an epistemic or doxastic logic. Note that it is disjoint from $\mathscr{L}$ in the sense that it cannot express objective formulas. Then we only keep the axioms K and D, and necessitation is modeled by an axiom saying that $\Box p$ is valid whenever p is a tautology of the propositional calculus. This system is called MEL (minimal epistemic logic) (Banerjee & Dubois, 2014).

The semantics of MEL is in terms of simple epistemic states $E \subseteq S$, and the satisfaction of $\Box p$ by E is then expressed by $E \subseteq [p]$, that is, p is true in all states that are not considered impossible for the agent, or alternatively, p is believed, that is, $N([p]) = 1$ for the Boolean necessity measure N equivalent to E. Indeed, axiom K ensures that $\Box(p \wedge q)$ is logically equivalent to $\Box p \wedge \Box q$, which is the axiom of necessity measures. Axiom D expresses that $N(A) = 1$ implies $\Pi(A) = 1$. It ensures consistency (models of MEL are then nonempty subsets of S).

In other words, the logic MEL bridges the gap between epistemic logic (dropping the idea of introspection) and the representation of belief by means of set functions, identifying the latter with modalities. Especially the modality $\Diamond p = \neg\Box\neg p$ corresponds to possibility measures.

3.2 The Logic of Graded Acceptance

The logic of acceptance MEL can be extended to graded necessity functions N using a simple multimodal logic. If L is a finite totally ordered scale of necessity degrees (understood as describing a gradation in accepted beliefs), we can expand the modal language of the MEL logic to allow for several belief modalities denoted by $\Box^\lambda$, for $\lambda \in L$ with $\lambda > 0$, where the sentence $\Box^\lambda p$ encodes the statement $N([p]) \geq \lambda$, and the $\Box$-modality in MEL corresponds to $\Box^1$ (expressing full belief). The language of this logic is thus a propositional language $\mathscr{L}^L_\Box$ where atoms are of the form $\Box^\lambda p$, $p \in \mathscr{L}$, for $\lambda \in L$ with $\lambda > 0$. Note that the formula $\neg\Box^\lambda p$ stands for $N([p]) < \lambda$, which, due to the assumed finiteness of L, can be expressed as $\Pi([\neg p]) \geq 1 - s(\lambda)$, which is encoded as $\Diamond^{s(\lambda)}\neg p$, for the next lower value $s(\lambda)$ to λ in the scale L.

The sublogic obtained by fixing the value λ is a copy of the logic MEL (it satisfies axiom K, axiom D, and necessitation). There is also the weakening axiom: $\Box^\lambda p \Rightarrow \Box^\mu p$ if $\mu \leq \lambda$, and axiom D is valid in the stronger form $\Box^\lambda p \Rightarrow \Diamond^1 p$. The semantics is in terms of L-valued possibility distributions π representing gradual epistemic states, and π satisfies $\Box^\lambda p$ if and only if $N([p]) \geq \lambda$, where N is based on π. The soundness and completeness of this logic, called GPL (generalized possibilistic logic), have been proved (Dubois, Prade, & Schockaert, 2017).

This logic is very expressive and enables one to reason about ignorance and defeasible beliefs. As shown in the previous reference, it can encode several nonmonotonic formalisms, especially:

- The older standard possibilistic logic (Dubois, Lang, & Prade, 1994) is obtained by restricting the language $\mathscr{L}^L_\Box$ to conjunctions of atomic epistemic statements $\Box^\lambda p$, which are written (p, λ) in the original syntax. This logic is nonmonotonic (Benferhat, Dubois, & Prade, 1998).
- Conditional logics with statements of relative belief of the form $N(p) > N(q)$ can be encoded by GPL formulas of the form $\bigvee_{\lambda \neq 0}(\Box^\lambda p \wedge \neg\Box^\lambda q)$.
- The System P of Kraus, Lehmann, and Magidor (1990) uses nonmonotonic conditional statements $p \mid\!\sim q$.

They express the plausible inference of q from p, which is modeled by the constraint $N(p \Rightarrow q) > N(p \Rightarrow \neg q)$.

- Answer-set programs can be expressed in GPL using a three-valued scale $L = \{1, \lambda, 0\}$, where $1 > \lambda > 0$, and requiring two necessity modalities: a strong one, $\Box^1$, and a weak one, $\Box^\lambda$.

In summary, GPL can be viewed as the logic of qualitative Baconian probabilities.

4. Nonadjunctive Settings for Rational Degrees of Belief

If we take it for granted that belief should come in degrees, and if, like Kyburg, we reject the adjunction rule, we are left with nonclassical logics where deduction is not closed under conjunction, like for instance the logic of risky knowledge (Kyburg & Teng, 2012). Or we give up the Boolean framework of logic altogether and concentrate on the properties of set functions that can model degrees of belief. The natural question is then whether degrees of belief should be additive at all. Fifty years ago, the answer was, Yes, of course. Since then, a number of proposals have emerged from which it follows that additivity should not be taken for granted. This section discusses reasons for questioning additive beliefs and focuses on some approaches to graded belief that are deliberately non-additive.

4.1 Can a Single Probability Distribution Capture Any Epistemic State?

The so-called Bayesian approach to subjective probability theory posits a uniqueness principle as a preamble to any kind of uncertainty modelling: Any state of knowledge is representable by a single probability distribution (see, for instance, Lindley, 1982). Note that indeed, if, following the fair bet procedure of de Finetti, an agent decides to directly assign subjective probabilities via buying prices to all possible outcomes in some game of chance, the coherence principle forces the agent to define a unique probability distribution.

Yet another mathematical attempt to justify probability theory as the only reasonable belief measure is the one by Richard T. Cox (1946). He relied on the Boolean structure of the set of events and on a number of postulates he considered compelling. Let $g(A \mid B) \in [0, 1]$ be a conditional degree of belief, A and B being events in a Boolean algebra, with $B \neq \emptyset$. The Cox axioms for conditional belief are as follows:

(i) $g(A \cap C \mid B) = F(g(A \mid C \cap B), g(C \mid B))$ (if $C \cap B \neq \emptyset$);

(ii) $g(A^c \mid B) = n(g(A \mid B))$ for $B \neq \emptyset$, where A^c is the complement of A in S;

(iii) The function F is supposed to be twice differentiable, with a continuous second derivative, while the function n is twice differentiable.

Cox claimed that, on such a basis, $g(A \mid B)$ must be isomorphic to a conditional probability measure.

This result has been repeated *ad nauseam* in the literature on artificial intelligence to justify probability combined with Bayes' rule as the only reasonable approach to represent and revise numerical degrees of belief (Horvitz, Heckerman, & Langlotz, 1986; Cheeseman, 1988; Jaynes, 2003). However, some reservations must be made. First, the original proof by Cox turned out to be faulty—see Paris (1994) for another proof, based on a weaker condition (iii) (it is enough that F be strictly monotonically increasing in each argument), but also on an additional technical density condition that requires an infinite setting. Moreover, Halpern (1999a,b) has shown that the result does not hold on finite spaces and that Cox's original conditions do not suffice to prove the result in the infinite setting. Independently of these technical issues, it should be noticed that postulate (i) sounds natural only if one takes the form of Bayes conditioning for granted; and postulate (ii) requires self-duality, which rules out representations of uncertainty due to partial ignorance, as seen later on. The above comments seriously weaken the alleged universality of Cox's results.

Applying the Bayesian credo as recalled above, justified via the avoidance of Dutch books, assuming fair prices for bets, or by obedience to Cox's axioms, forces the agent to use a single probability measure as the universal tool for representing uncertainty, whatever its source. This stance leads to serious representation difficulties already pointed out more than forty years ago (Shafer, 1976). For one, it means we give up making any difference between uncertainty due to incomplete information or ignorance, and uncertainty due to a purely random process, the next outcome of which cannot be predicted. One may indeed admit that additive degrees of beliefs are justified if they reflect extensive statistical evidence. But what if such information is not available?

Take the example of die tossing: The uniform probability assignment models the assumption that the die is fair. But if the agent assigns equal prices to bets assigned to all facets of the die, how can we interpret that? Is it because the agent is sure that the die is fair and its outcomes are driven by pure randomness (because, say, he or she could test it hundreds of times prior to placing the bets, or from counting cases)? Or is it because the agent who is given this die has just no idea whether the die is fair or not, and so has no reason to put more money on one facet than on another? Clearly the epistemic state of the agent is not the same in the first situation as in the second one. But the uniformly distributed probability function is mute about this issue and handles the two situations in the same way.

Next, the choice of a set of mutually exclusive outcomes depends on the chosen language, for example, the one used by the information source. However, several languages or points of view can co-exist for the same problem. Since there are several different possible representations of the state space, the probability assignment obtained from an agent will be language-dependent, especially in the case of ignorance: a uniform probability on one representation of the state space may conflict with a uniform one on another representation encoding of the same state space for the same problem, while in the case of ignorance this is the only available model. Shafer (1976) gives the following example: Consider the question of the existence of extraterrestrial life, about which the agent has no idea. If the variable v refers to the claim that life exists outside our planet (v = li), or not ($v = \neg$li), then the agent proposes $P_1(\text{li}) = P_1(\neg\text{li}) = ½$ on $S_1 = \{\text{li}, \neg\text{li}\}$. However, it makes sense to interpret "life" as the disjunction between animal life (ali), and vegetable life only (vli), which leads to the state space $S_2 = \{\text{ali, vli}, \neg\text{li}\}$. The ignorant Bayesian agent is then bound to propose $P_2(\text{ali}) = P_2(\text{vli}) = P_2(\neg\text{li}) = 1/3$. As "li" is the disjunction of "ali" and "vli," the distributions P_1 and P_2 are not compatible with each other, while they are both supposed to represent ignorance. Another example comes from noticing that expressing ignorance about a real-valued quantity by means of a uniform distribution for $x \in [a, b]$, a positive interval, is not compatible with a uniform distribution on $y = \log(x) \in [\log(a), \log(b)]$, while the agent has the same level of ignorance about x and y.

Finally, Ellsberg's (1961) paradox (see chapter 8.2 by Peterson, this handbook) showed quite early that, when expressing preferences between gambles consisting in drawing balls from an urn the content of which is ill known, many experiments have shown that people tend to systematically violate Savage's axioms (especially the sure-thing principle), because they are pessimistic about rewarding events of unknown probability. One way of accounting for the results of these experiments is to give up the additivity of degrees of belief.

The above limitations of the expressive power of single probability distributions have motivated the emergence of other approaches to representing uncertainty. Some of them, as seen above, give up the numerical setting of degrees of belief and use ordinal or qualitative structures, like qualitative possibility theory. Another option is to tolerate incomplete information in the probabilistic approach, which leads to different mathematical models of various levels of generality. These are reviewed in the rest of this chapter.

4.2 Shafer Belief Functions and the Merging of Uncertain Testimonies

The theory of evidence by Glenn Shafer (1976) can be viewed as a specific interpretation of Dempster's (1967) upper-and-lower-probability framework for handling imprecise statistical information, or as the revival of the concept of probability (as opposed to chance) invented in the 17th century by Jakob Bernoulli, and later by Johann Heinrich Lambert, dealing with the problem of representing and merging unreliable testimonies (Shafer, 1978). Shafer's book is clearly in the latter tradition.

The main issue is first to model unreliable testimonies. Suppose that a witness claims that proposition E is true but the receiver only partially believes this statement, considering that the witness is reliable with probability p. So p can be viewed as the degree of belief of the receiver in proposition E due to the unreliability of the witness. The additivity issue is raised by the question, to what proposition should the complementary weight $1-p$ be assigned? The regular probabilist would assign it to the negated proposition E^c. But if the testimony E is interpreted by the receiver as its negation, it means that the latter thinks this witness lies, that is, that he or she says E when knowing it to be false. There is another option, namely that due to the incompetence of the witness, there is a probability $1-p$ that the testimony is just useless. In the latter case, the probability $1-p$ is assigned not to E^c, but to the whole of S, that is, to the state of ignorance. That is to say, with probability p the receiver knows that E is true and nothing else, and with probability $1-p$ he/she knows nothing, which is modeled by a basic assignment m from the power set 2^S to $[0,1]$ such that $m(E) = p$ and $m(S) = 1-p$. A belief function that models such a simple unreliable testimony E is called a "simple support function."

More generally, consider a process whose outcomes are set-valued (i.e., imprecise) and uncertain (there is a probability value attached to each outcome). This is modeled by a more general basic assignment m from 2^S to [0, 1] such that $m(\emptyset) = 0$ and $\sum_{E \subseteq S} m(E) = 1$. Epistemic states E with $m(E) > 0$ are called "focal sets." The degree of belief $Bel(A)$ in a proposition A and its dual plausibility degree $Pl(A)$ are then defined by

$$Bel(A) = \sum\{m(E)\colon E \subseteq A,\ E \neq \emptyset\};$$
$$Pl(A) = \sum\{m(E)\colon E \cap A \neq \emptyset\}.$$

It is clear that the belief function *Bel* is non-additive, for example, $Bel(A) + Bel(A^c) \leq 1$, and the degree of plausibility is $Pl(A) = 1 - Bel(A^c)$. In the case of a simple support function, observe that when $A \neq S$, then $Bel(A) = p$ if $E \subseteq A$, and $Bel(A) = 0$ otherwise.

It is important to point out that belief functions generalize probabilities (recovered when for all E, $m(E) > 0$ implies that E is a singleton), Boolean necessity measures (recovered when $m(E) = 1$ for some epistemic state E), and also graded necessity measures (recovered when focal sets are all nested; an example of this is a simple support function). A degree of belief $Bel(A)$ clearly evaluates the probability of proving A from the available information. A plausibility degree $Pl(A)$ evaluates the probability that A is not logically incompatible with the available information. To use Cohen's (1977) terminology, belief functions join the probable and the provable, or bring Pascal and Bacon together. But doing so, belief functions are no longer additive, nor do they respect adjunction.

The major problem addressed by the 17th- and 18th-century pioneers is the merging of such testimonies. They proposed special cases of what is now known as Dempster's rule of combination: Let m_1 and m_2 be two mass assignments coming from independent sources. The result of the combination is a mass assignment $m_1 \otimes m_2$ defined by

for all $A \subseteq S$,
$(m_1 \otimes m_2)(A) = (1/K)\sum\{m_1(A_1) \cdot m_2(A_2)\colon A = A_1 \cap A_2\}$,
where $K = \sum\{m_1(A_1) \cdot m_2(A_2)\colon A_1 \cap A_2 \neq \varnothing\}$ and
$(m_1 \otimes m_2)(\varnothing) = 0$.

The assignment $m_1 \otimes m_2$ consists in intersecting any two overlapping focal sets, each coming from a distinct source, computing the probability of obtaining each subset A via such an intersection, and renormalizing the obtained mass assignment, as some pairs of focal sets may be conflicting. In the case of merging two simple support functions focusing on the same set E where $Bel_1(E) = m_1(E) = p_1$ and $Bel_2(E) = m_2(E) = p_2$, the resulting belief in E is $(Bel_1 \otimes Bel_2)(E) = p_1 + p_2 - p_1 p_2$, which leads to a reinforcement of the belief in E, a result already suggested by the pioneers of belief functions in the 17th and 18th century. This combination rule assumes that sources of information are independent, which makes the reinforcement effect plausible.

In Shafer's book, a major question was whether all belief functions can be expressed as the result of merging independent simple testimonies in the form of simple support functions. It turns out that only a subclass of belief functions, called *separable*, can be generated in this way. Later on, Smets (1995) tried to extend the notion of a simple support function so as to cover all belief functions. An extensive presentation of the theory of evidence as a theory of rational belief is proposed by Haenni (2009).

Finally, criteria for decision under uncertainty, where the latter is described by a belief function, are studied by Smets and Kennes (1994) and Jaffray (1989). The former propose to define a so-called *pignistic* probability measure from a belief function (generalizing Laplace's principle of insufficient reason), and apply the expected utility criterion. This probability measure coincides with the well-known Shapley value in game theory (Shapley, 1953) and in some sense projects the provable onto the probable. Jaffray (1989) proposes and axiomatizes an extension of Hurwicz's criterion.

4.3 Imprecise Probabilities, Desirability, and Generalized Betting Theory

The alternative approach to the modeling of degrees of belief consists in revisiting de Finetti's approach to subjective probability, dropping the constraint that the price proposed by a gambler for buying a lottery ticket should be fair. This view was pioneered by Cedric Smith (1961), Peter M. Williams (1975), and Robin Giles (1982), and more extensively developed by Peter Walley (1991).

In this approach, the agent offers buying prices for gambles. A gamble is a function f from S to the real line that expresses losses ($f(s) < 0$) or gains ($f(s) > 0$), depending on the actual state of affairs s. The gamble associated with a particular event is its characteristic function. The agent is not committed to selling such gambles at the same price as their buying price. The approach relies on so-called *desirable gambles* (Walley, 1991), which the agent would agree to buy for a positive price. The set of desirable gambles contains at least all positive gambles. Moreover, the sum of two desirable gambles is also considered desirable, and a desirable gamble remains desirable when multiplied by a positive constant. The maximal price at which the agent accepts to buy a gamble is the maximal value α such that $f - \alpha$ is desirable. It is called the *lower prevision* of a gamble f. It can be shown that given a set of gambles $\mathscr{G}$ and their lower previsions $LP(f)$, there is a convex set $\mathscr{P}_{\mathscr{G}}$ of probabilities, called a *credal set*, such that $LP(f) = \inf\{E_P(f)\colon P \in \mathscr{P}_{\mathscr{G}}\}$ is the lower expectation of f_i according to $\mathscr{P}_{\mathscr{G}}$, for all $f \in \mathscr{G}$, where $E_P(f)$ is the expectation of f with respect to probability P. One important point is that any convex set of probabilities can be represented by lower previsions on some family of gambles.

In this setting, the upper prevision $UP(f)$ of a gamble f is provably equal to $-LP(-f)$. The upper prevision $UP(f)$ is thus the minimal selling price of f. If the credal set attached to a set of gambles and its lower previsions is empty, then the proposal is inconsistent and the agent incurs a sure loss after buying and resolving these gambles. Avoiding sure loss means that $UP(f) \geq LP(f)$ for all gambles f.

Moreover, due to the interaction between gambles, it may be that the consistent buying prices proposed by the agent for gambles $f \in \mathcal{G}$ are too low and could be raised without altering the credal set. A set of buying prices $bp(f)$ for $f \in \mathcal{G}$ is said to be *coherent* if and only if $LP(f) = bp(f)$ for all $f \in \mathcal{G}$. In other words, a set of buying prices for a set of gambles $\mathcal{G}$ is coherent if and only if for any $f \in \mathcal{G}$,

$$\inf\{E_P(f): E_P(f) \geq bp(f) \text{ for all } f \in \mathcal{G}\} = bp(f).$$

Under this approach the degree of belief in proposition A is a coherent lower probability $P_*(A) = \inf\{P(A): P \in \mathrm{P}\} = LP(1_A)$, the lower prevision of its characteristic function, where P is the credal set induced by the lower prevision $LP(f)$ on some gambles.

Some remarks are in order to position this approach with respect to other approaches to rational degrees of belief:

- The epistemic state of the agent is here represented by a credal set, but there is no ill-known probability inside. In particular, the interval $[P_*(A), P^*(A)]$, where $P^*(A) = 1 - P_*(A^c)$, is not supposed to contain an ill-known subjective (nor an objective) probability of A. Just as for belief functions, degrees of belief are precise and modelled by coherent lower probabilities.
- Mathematically, belief functions are a special case of coherent lower probabilities. They are super-additive set functions at any order, while lower probabilities from any credal set only satisfy the inequality $P_*(A \cup B) \geq P_*(A) + P_*(B)$ when A and B are disjoint. In particular, the mass function recomputed from P_* instead of Bel in (3) (called the Moebius transform of P_*) exists, and is unique but not necessarily positive (Chateauneuf & Jaffray, 1989).
- An attempt to justify belief functions as the only rational approach to degrees of belief under a betting framework in the style of Walley (and not as unreliable testimonies) was recently published by Kerkvliet and Meester (2018).

The gamble approach leads to a decision rule that is specific to the imprecise-probability setting, namely, a gamble f is preferred to a gamble g if and only the gamble $h = f - g$ is desirable, that is, if the lower expectation of this gamble with respect to the corresponding credal set P is positive. This yields a partial ordering on gambles. It implies that $LP(f) \geq LP(g)$. The latter inequality is a decision rule that solves the Ellsberg paradox, in contrast with the decision rule $LP(f - g) \geq 0$ that satisfies Savage's sure-thing principle (see also chapter 8.2 by Peterson, this handbook).

4.4 Quantitative Possibility in the Setting of Imprecise Probability

It is natural to reconsider graded possibility and necessity measures in the setting of belief functions and imprecise probabilities. In fact they are at the crossroads of all non-additive approaches to uncertainty and may be interpreted in various ways:

- Necessity measures are a special case of belief functions. Their characteristic property is to have nested focal sets. In other words, they model coherent arguments in favor or disfavor of propositions. They are the only family of belief functions that obey the adjunction rule. Note that the weaker Baconian condition $\min(Bel(A), Bel(A^c)) = 0$ for all $A \subseteq S$ corresponds to overlapping (consistent) focal sets.
- As a consequence, necessity measures also stand for coherent lower probabilities. However, they correspond to a very cautious type of betting behavior, such that if the buying price for gambling on A is positive then the agent feels obliged to sell this gamble at the maximal price (Giles, 1982).
- One may borrow the operational semantics of the Bayesians to derive personal possibility and necessity degrees. If we adopt the framework of belief functions for representing an agent's knowledge and accept the idea that a belief function induces a pignistic probability for making decisions, then we may reverse this process. Given a subjective probability reflecting fair prices of gambles corresponding to random events, one may look for the least informative belief function that induces this subjective probability. It can be proved that it is always a necessity measure (Dubois, Prade, & Smets, 2008).
- Necessity functions induced on the unit interval by a suitable transformation of a Spohn ranking function (see chapter 5.3 by Kern-Isberner et al., this handbook) have nothing to do with lower probabilities. Basically, as shown in Spohn (1990), they are more closely related to powers of infinitesimal probabilities, for which the additivity axiom degenerates in the minitivity axiom.
- Yet another interpretation of possibility theory is in terms of likelihood. In statistical inference, given a parametric probabilistic model $P(\cdot \mid \theta)$ where $\theta \in \Theta$ is the parameter of the model, the probability $P(R \mid \theta)$

based on data set R is not the probability of θ based on R, only its likelihood. It represents a looser degree $lik(\theta) = P(R \mid \theta)$ of confidence in θ for the observer having received evidence R. Advocates of the likelihood approach (Edwards, 1972) refuse to attach prior probabilities to values of θ, basically because this quantity is not observable and is just a model artefact. Rather, it is natural to try and define the likelihood $lik(A)$ for any $A \subseteq \Theta$ from the values $lik(\theta)$ for $\theta \in A$. It has been shown that the only meaningful definition is $lik(A) = \max\{P(R \mid \theta)\colon \theta \in A\}$ (Dubois, Moral, & Prade, 1997; Dubois, 2006). Hence, in the absence of prior probabilities, a likelihood function can be interpreted as a possibility measure. However, this kind of possibility measure is determined only up to a multiplicative constant, a specific feature that makes likelihood theory yet another kind of possibility theory.

5. Conclusion

This chapter has presented a survey of various approaches to the notion of belief, reflecting the progress made in the last 50 years. It seems that the frontal opposition between degrees of belief and accepted beliefs, that is, the Pascalian and Baconian traditions, may be alleviated to some extent if we give up the requirement that degrees of belief should be additive. There is a range of mathematical models standing between probability and modal logics, some of which retain the adjunction rule of Baconian probabilities. Some approaches blend the two traditions and are consistent with the requirement that you cannot at the same time believe in a proposition and believe in its negation. The Baconian tradition also touches upon the issue of formal argumentation, on which there is an abundant literature today (see Haenni, 2009, for its connection with Shafer belief functions). Argumentation can be viewed as a rational approach to handle inconsistency in reasoning due to conflicting pieces of information. One may argue that Baconian probabilities (in the form of, e.g., necessity functions, ranking functions, and the like) represent imprecise but conflict-free information, while ordinary probabilities capture precise but conflicting observations. The new theories of belief deal with both imprecise and conflicting information and seem to bridge the gap between the two traditions of belief representation. One may then consider belief in a more dynamical setting, where starting with more or less probable conflicting evidence one proceeds towards the provable via a suitable deliberation process involving argumentation.

Notes

1. See the entry "Probabilité," accredited to Benjamin de Langes de Lubières (1714–1790), in *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers*, by D. Diderot and J.-B. Le Rond d'Alembert (http://enccre.academie-sciences.fr/encyclopedie/page/v13-p403).

2. Halpern (2003) calls set functions of this kind "plausibility measures," not even assuming a total order on events (replacing [0, 1] by a partially ordered set). However, this terminology may lead to confusion with Shafer's older plausibility functions (see section 4.2).

3. Or doxastic state—we do not distinguish between the two in this chapter.

4. Up to the presence or not of technical assumptions like well-ordering in the infinite setting.

References

Banerjee, M., & Dubois, D. (2014). A simple logic for reasoning about incomplete knowledge. *International Journal of Approximate Reasoning*, *55*, 639–653.

Benferhat, S., Dubois, D., & Prade, H. (1998). Practical handling of exception-tainted rules and independence information in possibilistic logic. *Applied Intelligence*, *9*, 101–127.

Benferhat, S., Dubois, D., & Prade, H. (1999). Possibilistic and standard probabilistic semantics of conditional knowledge bases. *Journal of Logic and Computation*, *9*, 873–895.

Chateauneuf, A., & Jaffray, J.-Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, *17*(3), 263–283.

Cheeseman, P. (1988). An inquiry into computer understanding (with comments and a rejoinder). *Computational Intelligence*, *4*, 58–142.

Choquet, G. (1953). Théorie des capacités [Theory of capacities]. *Annales de l'Institut Fourier*, *5*, 131–295.

Ciucci, D., & Dubois, D. (2013). A modal theorem-preserving translation of a class of three-valued logics of incomplete information. *Journal of Applied Non-Classical Logics*, *23*(4), 321–352.

Cohen, L. J. (1977). *The probable and the provable*. Oxford, England: Clarendon Press.

Cohen, L. J. (1980). Some historical remarks on the Baconian conception of probability. *Journal of the History of Ideas*, *41*, 219–231.

Cox, R. T. (1946). Probability, frequency, and reasonable expectation. *American Journal of Physics*, *14*, 1–13.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives [Foresight: its logical laws, its subjective sources], *Annales de l'Institut Poincaré*, *7*, 1–68. In H. E. Kyburg Jr. & H. E. Smokler (1964, Eds.), *Studies in subjective probability* (pp. 53–118), New York, NY: Wiley.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Statistics*, *28*, 325–339.

Dubois, D. (2006). Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis*, *51*(1), 47–69.

Dubois, D., Fargier, H., & Prade, H. (2004). Ordinal and probabilistic representations of acceptance. *Journal of Artificial Intelligence Research*, *22*, 23–56.

Dubois, D., Fargier, H., Prade, H., & Sabbadin, R. (2009). A survey of qualitative decision rules under uncertainty. In D. Bouyssou, D. Dubois, H. Prade, & M. Pirlot (Eds.), *Decision-making process: Concepts and methods* (pp. 435–473). London, England: ISTE & Wiley.

Dubois, D., Lang, J., & Prade, H. (1994). Possibilistic logic. In D. M. Gabbay, C. J. Hogger, J. A. Robinson, & D. Nute (Eds.), *Handbook of logic in artificial intelligence and logic programming: Vol. 3. Nonmonotonic reasoning and uncertain reasoning* (pp. 439–513). Oxford, England: Oxford University Press.

Dubois, D., Moral, S., & Prade, H. (1997). A semantics for possibility theory based on likelihoods. *Journal of Mathematical Analysis and Applications*, *205*, 359–380.

Dubois, D., & Prade, H. (1988). *Possibility theory: An approach to computerized processing of uncertainty*. New York, NY: Plenum Press.

Dubois, D., & Prade, H. (1998). Possibility theory: Qualitative and quantitative aspects. In D. M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems: Vol. 1. Quantified representation of uncertainty and imprecision* (pp. 169–226). Dordrecht, Netherlands: Kluwer.

Dubois, D., Prade, H., & Sabbadin, R. (2001). Decision-theoretic foundations of qualitative possibility theory. *European Journal of Operational Research*, *128*(3), 459–478.

Dubois, D., Prade, H., & Schockaert, S. (2017). Generalized possibilistic logic: Foundations and applications to qualitative reasoning about uncertainty. *Artificial Intelligence*, *252*, 139–174.

Dubois, D., Prade, H., & Smets, P. (2008). A definition of subjective possibility. *International Journal of Approximate Reasoning*, *48*, 352–364.

Edwards, A. W. F. (1972). *Likelihood*. Cambridge, England: Cambridge University Press.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*(4), 643–669.

Fargier, H., & Sabbadin, R. (2005). Qualitative decision under uncertainty: Back to expected utility. *Artificial Intelligence*, *164*(1–2), 245–280.

Giles, R. (1982). Foundations for a theory of possibility. In M. Gupta & E. Sanchez (Eds.), *Fuzzy information and decision processes* (pp. 183–195). Amsterdam, Netherlands: North-Holland.

Haenni, R. (2009). Non-additive degrees of belief. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (pp. 121–160). Berlin, Germany: Springer.

Halpern, J. Y. (1999a). A counterexample to theorems of Cox and Fine. *Journal of Artificial Intelligence Research*, *10*, 67–85.

Halpern, J. Y. (1999b). Technical addendum: Cox's theorem revisited. *Journal of Artificial Intelligence Research*, *11*, 429–435.

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

Hintikka, J. (1962). *Knowledge and belief*. Ithaca, NY: Cornell University Press.

Horvitz, E. J., Heckerman, D. E., & Langlotz, C. P. (1986). A framework for comparing alternative formalisms for plausible reasoning. In T. Kehler (Ed.), *Proceedings, AAAI-86: Fifth national conference on artificial intelligence, August 11–15, 1986* (Vol. 1, pp. 210–214). San Mateo, CA: Morgan Kaufmann.

Huber, F., & Schmidt-Petri, C. (Eds.). (2009). *Degrees of belief*. Berlin, Germany: Springer.

Jaffray, J. Y. (1989). Linear utility theory for belief functions. *Operations Research Letters*, *8*, 107–112.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge, England: Cambridge University Press.

Kerkvliet, T., & Meester, R. (2018). A behavioral interpretation of belief functions. *Journal of Theoretical Probability*, *31*, 2112–2128.

Koopman, B. O. (1940). The bases of probability. *Bulletin of the American Mathematical Society*, *46*, 763–774.

Kraus, S., Lehmann, D., & Magidor, M. (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, *44*, 167–207.

Kyburg, H. (1961). *Probability and the logic of rational belief*. Middletown, CT: Wesleyan University Press.

Kyburg, H. (1997). The rule of adjunction and reasonable inference. *Journal of Philosophy*, *94*, 109–125.

Kyburg, H., & Teng, C. M. (2012). The logic of risky knowledge, reprised. *International Journal of Approximate Reasoning*, *53*, 274–285.

Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, *123*(2), 131–171.

Lewis, D. K. (1973). *Counterfactuals*. London, England: Basil Blackwell (2nd ed.). Worcester, England: Billing.

Lindley, D. V. (1982). Scoring rules and the inevitability of probability. *International Statistics Review*, *50*, 1–26.

Paris, J. B. (1994). *The uncertain reasoner's companion: A mathematical perspective*. Cambridge, England: Cambridge University Press.

Ramsey, F. P. (1931). Truth and probability. In F. P. Ramsey, *The foundations of mathematics and other logical essays* (R. B. Braithwaite, Ed., pp. 156–198). London, England: Kegan, Paul, Trench, Trubner. (Original work 1926)

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Shackle, G. L. S. (1949). *Expectation in economics*. Cambridge, England: Cambridge University Press.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.

Shafer, G. (1978). Non-additive probabilities in the work of Bernoulli and Lambert. *Archive for History of Exact Sciences*, *19*(4), 309–370.

Shapley, L. S. (1953). A value for *n*-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games* (Vol. II, pp. 307–317). Princeton, NJ: Princeton University Press.

Smets, P. (1995). The canonical decomposition of a weighted belief. In C. S. Mellish (Ed.), *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (pp. 1896–1901). San Mateo, CA: Morgan Kaufmann.

Smets, P., & Kennes, R. (1994). The transferable belief model. *Artificial Intelligence*, *66*, 191–234.

Smith, C. A. B. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society B*, *23*, 1–37.

Snow, P. (1999). Diverse confidence levels in a probabilistic semantics for conditional logics. *Artificial Intelligence*, *113*, 269–279.

Spohn, W. (1988). Ordinal conditional functions: A dynamic theory of epistemic states. In W. L. Harper & B. Skyrms (Eds.), *Causation in decision, belief change, and statistics* (Vol. 2, pp. 105–134). Dordrecht, Netherlands: Kluwer.

Spohn, W. (1990). A general, nonprobabilistic theory of inductive reasoning. In R. D. Shachter, T. S. Levitt, L. N. Kanal, & J. F. Lemmer (Eds.), *Uncertainty in artificial intelligence* (Vol. 4, pp. 149–158). Amsterdam, Netherlands: North Holland.

Sugeno, M. (1977). Fuzzy measures and fuzzy integrals—A survey. In M. M. Gupta, G. N. Saridis, & B. R. Gaines (Eds.), *Fuzzy automata and decision processes* (pp. 89–102). Amsterdam, Netherlands: North Holland.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities* (Monographs on Applied Statistics and Probability 42). New York, NY: Chapman and Hall.

Williams, P. M. (2007). Notes on conditional previsions. *International Journal of Approximate Reasoning*, *44*, 366–383 (unpublished research report, 1975).

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, *1*, 3–28.

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 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