International Journal of Approximate Reasoning 123 (2020) 26–40

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Depth-bounded Belief functions ☆

Paolo Baldi *, Hykel Hosni

Department of Philosophy, University of Milan, Italy

ARTICLE INFO

Article history:
Received 30 September 2019
Received in revised form 12 March 2020
Accepted 13 May 2020
Available online 22 May 2020

Keywords:
Belief functions
Uncertain reasoning
Depth-bounded logics
Probability

ABSTRACT

This paper introduces and investigates Depth-bounded Belief functions, a logic-based representation of quantified uncertainty. Depth-bounded Belief functions are based on the framework of Depth-bounded Boolean logics [4], which provide a hierarchy of approximations to classical logic. Similarly, Depth-bounded Belief functions give rise to a hierarchy of increasingly tighter lower and upper bounds over classical measures of uncertainty. This has the rather welcome consequence that "higher logical abilities" lead to sharper uncertainty quantification. In particular, our main results identify the conditions under which Dempster-Shafer Belief functions and probability functions can be represented as a limit of a suitable sequence of Depth-bounded Belief functions.

1. Introduction and motivation

The link between the rules of rational reasoning and the probabilistic representation of uncertainty is a strong and old one. Since Jacob Bernoulli's 1713 *Ars Conjectandi*, a number of arguments have been put forward to the effect that departing from a probabilistic assessment of uncertainty leads to *irrational* patterns of behaviour, as fixed by the well known results of de Finetti and Savage [7,25] ([16] and [22] provide recent introductory reviews).

Over the past few decades, however, a number of concerns have been raised against the adequacy of probability as a norm of rational reasoning and decision-making. Following the lead of [9], whom in turn found himself on the footsteps of [15] and [14], many decision theorists took issue with the normative adequacy of probability. As a result, considerable formal and conceptual effort has gone into extending the scope of the probabilistic representation of uncertainty, as briefly recalled in Section 1.1 below.

One key commonality among those "non probabilistic" approaches is the conviction that probability fails on representational grounds. On those grounds, they insist that the rational representation of uncertainty need not necessitate that all uncertainty be quantified probabilistically. But this was precisely a key concern tackled by the "Mathematical theory of evidence" put forward by Glenn Shafer and which has become known as the Dempster-Shafer theory of Belief functions (see [8] for a retrospective on the development of the theory and a comprehensive bibliography). Shafer's original aim in [27] was to provide a general theory of evidence and uncertain reasoning: in his interpretation, given a Belief function Bel, the value $Bel(\theta)$ stands for the degree of support that a piece of evidence provides for the event or proposition θ, according to the judgment of a certain agent. Since then, various interpretations of Belief functions have been advanced, especially concerning their relation with probability. Of particular relevance for our present purposes is Shafer's own later

☆ This research was funded by the Department of Philosophy "Piero Martinetti" of the University of Milan under the Project "Departments of Excellence 2018-2022" awarded by the Ministry of Education, University and Research (MIUR).

* Corresponding author.
E-mail addresses: paolo.baldi@unimi.it (P. Baldi), hykel.hosni@unimi.it (H. Hosni).

https://doi.org/10.1016/j.ijar.2020.05.001
0888-613X/

work [28], in which he suggests to take degrees of belief to be degrees of support to answers for a given question, for which no probability distribution is known, but that nevertheless can be computed by resorting to answers to a related question, which imply the original ones, and for which at the same time a probability distribution is actually known. This approach is exemplified by assessing the degree of support for a certain proposition in terms of reliability of witnesses (which can be probabilistically evaluated) asserting that proposition. An instance of such view of Belief functions, relevant for us later, is also the so-called *probability of provability* interpretation, as presented e.g. in [23] and [19].

Inspired by the setting of [19], the investigation reported in this paper tackles the representational shortcomings in the probabilistic uncertain reasoning from a *logical* point of view. We observe that these issues can be traced back to analogous shortcomings of classical logic, which provides a decidedly poor representation of the key notion of "information", cognate concept of "evidence". Armed with this simple but very important observation, we take a step back and *rethink the logic* in the first place. Hence, developing ideas first discussed in [5], we will cast the question of quantifying rational degrees of belief in the framework of Depth-bounded Boolean logics, recalled in Section 2 below.

This approach leads to a definition of *Depth-bounded Belief functions*, which are introduced in Section 3. On this basis, in Section 4 we provide a novel representation of classical Belief functions, which we see as the limit of sequences of Depth-bounded Belief functions. We will then investigate, in Section 5 and 6 interesting subclasses of Depth-bounded Belief functions, which arise by some principled restriction of the set of formulas where rational degrees of belief are quantified. In particular, one such subclass lead also to a representation theorem for classical Probability functions, which are obtained as the limit of sequences of (suitably constrained) Depth-bounded Belief functions.

Before delving into the details of our proposal, however, it will be useful to motivate its relevance against the wider landscape of rational reasoning and decision-making under uncertainty.

1.1. Uncertainty, ignorance and information

Uncertainty has to do, of course, with not knowing, and in particular not knowing the outcome(s) of an event of interest, or the value of a random variable. Ignorance has more subtle features, and is often thought of as our inability to quantify our own uncertainty. In [15], Knight gave this impalpable distinction an operational meaning in actuarial terms. He suggested that the presence of ignorance is detected by the absence of a complete insurance market for the goods at hand. On the contrary, a complete insurance market provides an operational definition of *probabilistically quantifiable* uncertainty. Contemporary followers of Knight insist that the inevitability of ignorance implies that not all uncertainty is probabilistically quantifiable and seek to introduce more general norms of rational belief and decision under "Knightian uncertainty" or "ambiguity" (see e.g. [12]). A telling illustration of this argument from information is due to David Schmeidler [26]:

> The probability attached to an uncertain event does not reflect the heuristic amount of information that led to the assignment of that probability. For example, when the information on the occurrence of two events is symmetric they are assigned equal probabilities. If the events are complementary the probabilities will be 1/2 independent of whether the symmetric information is meager or abundant.

Gilboa [12] interprets Schmeidler's observation as expressing a form of "cognitive unease", namely a feeling that the theory of subjective probability which springs naturally from Bayesian epistemology is silent on one fundamental aspect of rationality, namely how the available information, or the lack thereof, guides the process of uncertainty quantification. But why is it so? Suppose that some matter is to be decided by the toss of a coin. According to Schmeidler's line of argument, I should prefer tossing my own, rather than some one else's coin, on the basis, say of the fact that I have never observed signs of "unfairness" in my coin, whilst I just don't know anything about the stranger's coin. This qualitative information should be reflected in how uncertainty is to be rationally evaluated.

Similar considerations had been put forward in the foundations of statistics and later penetrated the broad field of uncertainty in Artificial Intelligence. As anticipated above, an early amendment of probability theory aimed at capturing the asymmetry between uncertainty and ignorance is the theory of Belief functions. Key to representing this asymmetry is the relaxation of the additivity axiom of probability. This in turn may lead to situations in which the *probabilistic excluded middle* does not hold. That is to say an agent could rationally assign belief less than 1 to the classical tautology $\phi \vee \neg\phi$. Indeed, as we now illustrate, the problem with normalising on the tautologies of *classical* logic can be taken as a starting point for more general considerations.

1.2. Probability and classical logic

It is well-known that every probability function arises from distributing the unit mass across the 2^n atoms of the Boolean (Lindenbaum) Algebra generated by the propositional variables $\{p_1, \dots p_n\}$ of a language $\mathcal{L}$, and conversely, that a probability function on formulas over $\mathcal{L}$ is completely determined by the values it takes on such atoms (see, e.g. [19] for a presentation in the spirit of our work). Such a representation makes explicit the twofold role played by classical logic and its extensions in the theory of probability.

First, it provides a language in which events – the bearers of probability – can be expressed, combined and evaluated. The precise details depend on the framework. See [10] for a characterisation of probability on classical logic, and [11] for the general case of Dempster-Shafer Belief functions on the many-valued extension of classical logic. In contrast, the measure-theoretic presentations of probability identifies events with subsets of the field generated by a given sample space Ω. A popular interpretation for Ω is that of the elementary outcomes of some experiment, a view endorsed by A.N. Kolmogorov, who insisted on the generality of his axiomatisation. More precisely, let $\mathcal{M} = (\Omega, \mathcal{F}, P)$ be a *measure space* where, $\Omega = \{\omega_1, \omega_2 \ldots\}$ is the set of elementary outcomes, $\mathcal{F} = 2^\Omega$ is the field of sets ($\sigma-$algebra) over Ω. We call *events* the elements of $\mathcal{F}$, and $P : \mathcal{F} \to [0, 1]$ a *probability measure* if it is normalised, monotone and σ-additive, i.e.

(K1) $P(\Omega) = 1$
(K2) $A \subseteq B \Rightarrow P(A) \leq P(B)$
(K3) If $\{E\}_i$ is a countable family of pairwise disjoint events in $\mathcal{F}$ then $P(\bigcup_i E_i) = \sum_i P(E_i)$

The Stone representation theorem for Boolean algebras and the representation of probability functions recalled above guarantee that the measure-theoretic axiomatisation of probability is equivalent to the logical one yielded by classical logic, which is obtained by letting a function from the formulas $Fm_{\mathcal{L}}$ of a language $\mathcal{L}$ to the real unit interval be a *probability function* if

(PL1) $\models \theta \Rightarrow P(\theta) = 1$
(PL2) $\models \neg(\theta \wedge \phi) \Rightarrow P(\theta \vee \phi) = P(\theta) + P(\phi)$.

Less straightforward but equally compelling is the case yielded by the many-valued extension of classical logic, see [17].

Obvious as the logical "translation" of the Kolmogorov axioms may be, it highlights a second and crucial role for logic in the theory of probability, which is best appreciated by focussing on the consequence relation $\models$.

In its measure-theoretic version, the normalisation axiom (K1) is quite uncontroversial. Less so, if framed in terms of classical tautologies, as in PL1. Indeed the arguments against the probabilistic representation of rational belief recalled above, emerge now formally. For $\models$ interprets symmetrically "knowledge" and "ignorance" as captured by the fact that $\models \theta \vee \neg\theta$ is a tautology. Indeed similarly bothersome consequences follow directly from $PL1$ and $PL2$, namely

1. $P(\neg\theta) = 1 - P(\theta)$
2. $\theta \models \phi \Rightarrow P(\theta) \leq P(\phi)$

Those examples suffice to make a key point: Many of the features of probability with which the critics recalled above take issue clearly have their logical roots in the *semantics of classical logic*.

The logical framing of probability allows us to refine this analysis. For it is the semantics of classical logic that provides the uncertainty resolution device for the evaluation of probability. This is best illustrated by the piecemeal identification of "events" with the "sentences" of the logic. On the one hand, an *event*, understood classically, either happens or not. A sentence expressing an event, on the other hand is evaluated in the binary set as follows

$$v(\theta) = \begin{cases} 1 \text{ if the event obtained} \\ 0 \text{ otherwise.} \end{cases}$$

Hence, the probability of an event $P(\theta) \in [0, 1]$ measures the agent's degree of belief that the event did or will obtain. Finding this out is, in most applications, relatively obvious. However, as pointed out in [10], a general theory of what it means for "states of the world" to "resolve uncertainty" is far from trivial.

A methodologically more adequate way of evaluating events arises by taking an *information-based* view on uncertainty resolution. The key difference with the previous, classical case, lies in the fact that this leads naturally to a *partial* evaluation of events, that is

$$v^i(\theta) = \begin{cases} 1 \text{ if I am informed that } \theta \\ 0 \text{ if I am informed that } \neg\theta \\ * \text{ if I am not informed about } \theta. \end{cases}$$

Quite obviously standard probability logic does *not* apply here, because the classical resolution of uncertainty has no way of expressing the $*$ condition. To model this we turn to the theory of Depth-bounded Boolean Logics [4,3].

2. An informational view of propositional logic: Depth-bounded Boolean logics

Depth-bounded Boolean logics are based on the idea that classical propositional connectives should be given an informational meaning. This is achieved by replacing the notions of "truth" and "falsity" by "informational truth" and "informational falsity", namely *holding the information* that a sentence φ is true, respectively false. Here, by saying that an agent *a holds* the information that φ is true or false it is meant that this information is *available* to *a* in the sense that *a* is ready to *act upon it*.

$\wedge$	1	0	$*$
1	1	0	$*$
0	0	0	0
$*$	$*$	0	$*, 0$

$\vee$	1	0	$*$
1	1	1	1
0	1	0	$*$
$*$	1	$*$	$*, 1$

$\neg$	
1	0
0	1
$*$	$*$

Fig. 1. Informational tables for the classical operators.

Both proof-theoretic and model-theoretic presentations of Depth-bounded Boolean logics are available, and indeed owing to completeness results, they are provably equivalent. Whilst our main results below are cast proof-theoretically, it is certainly helpful to the unfamiliar reader if we start by presenting Depth-bounded Boolean logics semantically. Indeed, as anticipated, they arise naturally by shifting the interpretation of boolean tables from "truth" to "informational truth". This apparently innocent shift makes undesirable the classical principle of Bivalence: it may well be that for a given φ, we neither hold the information that φ is true, nor do we hold the information that φ is false. So, the informational semantics which we are now ready to recall, has a built-in feature that marks the asymmetry between "knowledge" and "ignorance".

2.1. Semantics

For the rest of the paper we fix a language $\mathcal{L}$, over a finite set $Var = \{p_1, \dots, p_n\}$ of propositional variables. We let $Fm_{\mathcal{L}}$ be the formulas built from the propositional variables by the usual classical connectives $\wedge, \vee, \neg, \rightarrow$ and the constant for falsum $\curlywedge$. For each $p_i \in Var$ we denote by $\pm p_i$ any of the literals p_i and $\neg p_i$. Finally, for each set of formulas Γ we denote by $Sf(\Gamma)$ the subformulas of the formulas in Γ. We use the values 1 and 0 to represent, respectively, informational truth and falsity. When a sentence takes neither of these two defined values, we say that it is *informationally indeterminate*. It is technically convenient to treat informational indeterminacy as a third value that we denote by "$*$".[1] The three values are partially ordered by the relation $\preceq$ such that $v \preceq w$ ("v is less defined than, or equal to, w") if, and only if, $v = *$ or $v = w$ for $v, w \in \{0, 1, *\}$.

Note that the old familiar boolean tables for $\wedge, \vee$ and $\neg$ are still intuitively sound under this informational reinterpretation of 1 and 0. However, they are no longer exhaustive: they do not tell us what happens when one or all of the immediate constituents of a complex sentence take the value $*$. A remarkable consequence of this approach is that the semantics of $\vee$ and $\wedge$ becomes, as first noticed by Quine [24], *non-deterministic*. In some cases an agent a may accept a disjunction $\varphi \vee \psi$ as true while abstaining on both components φ and ψ. This is often the case when, trying to log in to a website, we are prompted with the error message "either your username (φ) or password (ψ) are wrong". Possessing this disjunctive piece of information does not give us any definite information about either disjunct, and therefore its seems only rational to refrain to act as if either was true or false. Similarly, a may reject a conjunction $\varphi \wedge \psi$ as false while abstaining on both components. Suppose a holds the information that Alice and Bob are not siblings $\neg(\varphi \wedge \psi)$. With this information a is clearly not in a position to either assent or dissent to sentences like "Is John Alice's father?" and "Is John Bob's father?". However a should certainly dissent to "John is Alice's father and Bob's father". Continuing with informal examples of this sort, one can also see that depending on the information actually possessed by the agent, when φ and ψ are both assigned the value $*$, the disjunction $\varphi \vee \psi$ may take the value 1 or $*$, and the conjunction $\varphi \wedge \psi$ may take the value 0 or $*$. This motivates the informational tables introduced in Fig. 1.

As a consequence of this informational interpretation, the classical boolean tables for the $\vee, \wedge$ and $\neg$ should be replaced by the "informational tables" in Fig. 1, where the value of a complex sentence, in some cases, is not uniquely determined by the value of its immediate components. A non-deterministic table for the informational meaning of the Boolean conditional can be obtained in the obvious way, by considering $\varphi \rightarrow \psi$ as having the same meaning as $\neg\varphi \vee \psi$ [see3, p. 82].

2.2. Derivation

The inferences which are allowed by taking the closure of the informational tables just described can be characterised equivalently in terms of Introduction (Table 1) and Elimination rules (Table 2).

The rules in Tables 1 and 2 determine a notion of 0-depth consequence relation as follows.

Definition 1. For any set of formulas $\Gamma \cup \{\alpha\} \subseteq Fm_{\mathcal{L}}$, we let $\Gamma \vdash_0 \alpha$ iff there is a sequence of formulas $\alpha_1, \dots, \alpha_n$ such that $\alpha_n = \alpha$ and each formula α_i is either in Γ or obtained by application of the rules in Table 1 and Table 2 on the formulas α_j with $j < i$.

The key feature of the consequence relation $\vdash_0$ is that only information actually possessed by the agent is allowed in a "0-depth deduction". This makes a sharp distinction with classical deduction where unbounded use can be made of *virtual information*, i.e. information not actually possessed by the agent at the time of carrying out the deduction. Virtual information is used to fill out, in all possible completions, gaps in the agent's information, as illustrated by Example 1 below.

[1] This is the symbol for "undefined", the bottom element of the information ordering, not to be confused with the "falsum" logical constant.

Table 1
Introduction rules.

$\dfrac{\varphi \quad \psi}{\varphi \wedge \psi}(\wedge\mathcal{I})$	$\dfrac{\neg\varphi}{\neg(\varphi \wedge \psi)}(\neg\wedge\mathcal{I}1)$	$\dfrac{\neg\psi}{\neg(\varphi \wedge \psi)}(\neg\wedge\mathcal{I}2)$
$\dfrac{\neg\varphi \quad \neg\psi}{\neg(\varphi \vee \psi)}(\neg\vee\mathcal{I})$	$\dfrac{\varphi}{\varphi \vee \psi}(\vee\mathcal{I}1)$	$\dfrac{\psi}{\varphi \vee \psi}(\vee\mathcal{I}2)$
$\dfrac{\varphi \quad \neg\psi}{\neg(\varphi \to \psi)}(\neg\to\mathcal{I})$	$\dfrac{\neg\varphi}{\varphi \to \psi}(\to\mathcal{I}1)$	$\dfrac{\psi}{\varphi \to \psi}(\to\mathcal{I}2)$
$\dfrac{\varphi \quad \neg\varphi}{\curlywedge}(\curlywedge\mathcal{I})$	$\dfrac{\varphi}{\neg\neg\varphi}(\neg\neg\mathcal{I})$	

Table 2
Elimination rules.

$\dfrac{\varphi \vee \psi \quad \neg\varphi}{\psi}(\vee\mathcal{E}1)$	$\dfrac{\varphi \vee \psi \quad \neg\psi}{\varphi}(\vee\mathcal{E}2)$	$\dfrac{\neg(\varphi \vee \psi)}{\neg\varphi}(\neg\vee\mathcal{E}1)$
$\dfrac{\neg(\varphi \vee \psi)}{\neg\psi}(\neg\vee\mathcal{E}2)$	$\dfrac{\varphi \wedge \psi}{\varphi}(\wedge\mathcal{E}1)$	$\dfrac{\varphi \wedge \psi}{\psi}(\wedge\mathcal{E}2)$
$\dfrac{\neg(\varphi \wedge \psi) \quad \varphi}{\neg\psi}(\neg\wedge\mathcal{E}1)$	$\dfrac{\neg(\varphi \wedge \psi) \quad \psi}{\neg\varphi}(\neg\wedge\mathcal{E}2)$	$\dfrac{\varphi \to \psi \quad \varphi}{\psi}(\to\mathcal{E}1)$
$\dfrac{\varphi \to \psi \quad \neg\psi}{\varphi}(\to\mathcal{E}2)$	$\dfrac{\neg(\varphi \to \psi)}{\varphi}(\neg\to\mathcal{E}1)$	$\dfrac{\neg(\varphi \to \psi)}{\neg\psi}(\neg\to\mathcal{E}2)$
$\dfrac{\neg\neg\varphi}{\varphi}(\neg\neg\mathcal{E})$	$\dfrac{\curlywedge}{\varphi}(\curlywedge\mathcal{E})$	

While $\vdash_0$ allows for no use of virtual information, the central idea of Depth-bounded Boolean Logics consists in keeping track of the amount k of virtual information which agents are allowed to use in their deductions. This naturally leads to the recursive definition of the consequence relation $\vdash_k$, for $k > 0$, as follows.

Definition 2. For each $k > 0$ and set of formulas $\Gamma \cup \{\alpha\} \subseteq Fm_{\mathcal{L}}$, we let $\Gamma \vdash_k \alpha$ iff there is a finite sequence of formulas $\alpha_1, \ldots, \alpha_n$ and a formula $\beta \in Sf(\Gamma \cup \{\alpha\})$, such that $\alpha_n = \alpha$, and for each α_i, with $i < n$, we have $\alpha_1, \ldots, \alpha_{i-1}, \beta \vdash_{k-1} \alpha_i$ and $\alpha_1, \ldots, \alpha_{i-1}, \neg\beta \vdash_{k-1} \alpha_i$.

In other words, we suppose that β is a piece of "virtual information" which is not actually possessed by the agent at level $k-1$ but which can be seen to be sufficient to derive α through case-based reasoning.[2] Note that the consequence relation presented in Definition 2 is an instance of a *strong* depth-bounded consequence relations, which is transitive, in contrast to its *weak* counterpart, see [4] for an exhaustive discussion. Variants of both strong and weak depth-bounded consequence relations arise when allowing different kind of formulas to be used as virtual information (*virtual spaces* in the terminology of [4]).

Example 1. Consider the excluded middle formula $p \vee \neg p$. Direct inspection of the rules in Tables 1 and 2 shows that the formula is not 0-depth derivable, i.e. $\nvdash_0 p \vee \neg p$. However, if we allow the use of virtual information p, we find out that both $p \vdash_0 p \vee \neg p$ and $\neg p \vdash_0 p \vee \neg p$. From this follows that $\vdash_1 p \vee \neg p$.

Let us just recall a small set of properties of Depth-bounded Boolean logics, which will play a role in what follows.

First, it is shown in [4] that the relation $\vdash_0$ is sound and complete with respect to the consequence relation defined on the basis of the non-deterministic semantics illustrated in Section 2.1. As an immediate consequence we can observe that Example 1 implies that it is not the case that $\phi \vee \neg\phi$ is a 0-depth tautology. This matches quite naturally the concerns raised against the so-called probabilistic excluded middle. In the framework of Depth-bounded Boolean logics, if an agent has no information about ϕ, then they should not be forced to assign it the highest degree of belief, for it is not a tautology.

Second, the recursive definition of $\vdash_k$ ensures that the bounded use of virtual information is monotonic and eventually becomes "unbounded", thereby providing a hierarchy of consequence relations approximating the classical one.

Theorem 1. *[4] The relations $\vdash_k$ approximate the classical consequence relation $\vdash$, that is,* $\lim_{k\to\infty} \vdash_k = \vdash$.

[2] The definition of $\vdash_k$ can be reformulated in terms of calculi that add to the rules in Table 1 and 2, a *branching rule*, to be applied only in limited form. Such rule (see e.g. [6]) bears some similarity to the elimination rule for disjunction in natural deduction, where the virtual information plays the role of formulas to be discharged.

Finally, as we shall discuss in the concluding section of this paper, the hierarchy of depth-bounded logics has a very nice computational feature: each consequence relation $\vdash_k$ is polynomial.

3. Belief functions and Depth-bounded logic

In this section we will use the tools of Depth-bounded logics and the informational perspective to rethink the very idea of degrees of belief, making it sensitive to the distinction between the manipulation of actual and virtual information [5].

First, for each formula $\alpha \in Fm_{\mathcal{L}}$, we define the function $v^k_\alpha : Fm_{\mathcal{L}} \to \{0,1\}$ such that $v^k_\alpha(\varphi) = 1$ iff $\alpha \vdash_k \varphi$ and $v^k_\alpha(\varphi) = 0$ otherwise.

We add to these the function v^k_* (where $*$ is a symbol not occurring in $Fm_{\mathcal{L}}$), such that $v^k_*(\varphi) = 1$ iff $\vdash_k \varphi$. Note that, since the 0-depth logic $\vdash_0$ does not have tautologies [4], $v^0_*(\varphi) = 0$ for each $\varphi \in Fm_{\mathcal{L}}$. We will also have a valuation $v^k_\curlywedge$, assigning $v^k_\curlywedge(\alpha) = 1$ iff $\curlywedge \vdash_k \alpha$. In the following, we call all the functions v^k_α, for $\alpha \in Fm_{\mathcal{L}} \cup \{*\}$ the *k-depth information states*.

A couple of observations are in order. First, the evaluations v^k_α can be thought of as lower approximations of a set of partial evaluations satisfying α, which send all the undetermined truth values to 0. An upper approximation, sending all the undetermined truth values to 1 is just obtained by letting $w^k_\alpha(\varphi) = 0$ iff $\alpha \vdash_k \neg\varphi$ and $w^k_\alpha(\varphi) = 1$ otherwise. Clearly, $w^k_a(\varphi) = 1 - v^k_\alpha(\neg\varphi)$. We can relate the pair (v^k_α, w^k_α) to the nondeterministic semantics (the indeterminate value should correspond to the case where $v^k_\alpha(\varphi) \neq w^k_\alpha(\varphi)$).

Finally, note that the function associating to each formula α the corresponding v^k_α is not injective: two distinct formulas α and β might determine the same function v^k_α and v^k_β. We can thus think of such functions as determining a partition of the set of formulas $Fm_{\mathcal{L}}$.

The first step in our characterisation of Belief functions based on Depth-bounded Boolean logics consists in defining a basic assignment over the set of formulas $Fm_{\mathcal{L}}$. In doing so, we assume that it is nonzero over a finite subset of $Fm_{\mathcal{L}}$. In addition to being mathematically convenient, this matches our general concern for maintaining the asymmetry between knowledge and ignorance. For, contrary to what happens for probability, assigning a degree 0 to the evidence for a formula α is not equivalent, in our setting, to assigning 1 to the negation $\neg\alpha$. Assignment of degree 0, even to infinitely many formulas, come in a sense at no cost.

Definition 3 *(k-depth mass function).* Let $m_k : Fm_{\mathcal{L}} \cup \{*\} \to [0,1]$ be such that

$$Supp(m_k) = \{\alpha \in Fm_{\mathcal{L}} \cup \{*\} \mid m_k(\alpha) \neq 0\}$$

is finite. Then m_k is a *k-depth mass function* if

1. $\sum_{\alpha \in Supp(m_k)} m_k(\alpha) = 1$
2. $\alpha \notin Supp(m_k)$ if $v^k_\alpha = v^k_\curlywedge$

The idea is that $m_k(\alpha)$ expresses the portion of belief that a k-depth bounded agent would assign *exclusively* to α, based on its actual information. On the other hand, $m_k(*)$ stands just for the portion of belief not assigned to any proposition. In case there is a formula γ, such that $\alpha \vdash_0 \gamma$ for each $\alpha \in Supp(m_k)$, we will denote the support by $Supp_\gamma(m_k)$.

Mass functions m_k whose supports have the latter form, express the more realistic situation where an agent is judging the evidence, when already in possession of a background information, represented by γ. Henceforth, if no such formula exist, with a slight abuse of notation, we will sometimes denote the support by $Supp_*(m_k)$, so that we can uniformly use the notation $Supp_\gamma$, for γ in $Fm_{\mathcal{L}} \cup \{*\}$. Adapting from the terminology in use for classical Belief functions, we will call the formulas in $Supp_\gamma(m_k)$ *focal* formulas.

We are now ready to introduce the notion of k-depth Belief function.

Definition 4 *(k-depth Belief function).* Let $\gamma \in Fm_{\mathcal{L}} \cup \{*\}$, and let m_k be a k-depth mass function with support $Supp_\gamma$. We define a corresponding *k-depth Belief function* as follows:

$$B_k(\varphi|\gamma) := \sum_{\alpha \in Supp_\gamma(m_k)} m_k(\alpha) \cdot v^k_\alpha(\varphi)$$

for any $\varphi \in Fm_{\mathcal{L}}$. Correspondingly, a *k-depth plausibility function* is

$$Pl_k(\varphi|\gamma) := \sum_{\alpha \in Supp_\gamma(m_k)} m_k(\alpha) \cdot w^k_\alpha(\varphi)$$

where we take into account all the α compatible with φ i.e. those which do not prove $\neg\varphi$.

It is easy to see that

$$Pl_k(\varphi|\gamma) = 1 - B_k(\neg\varphi|\gamma).$$

Given a formula φ and a background information γ, the k-depth belief an agent possesses about φ can be framed in terms of the interval $[B_k(\varphi|\gamma), Pl_k(\varphi|\gamma)]$.

Note the strong link between the representation of uncertainty provided by the interval $[B_k(\varphi|\gamma), Pl_k(\varphi|\gamma)]$ and the nondeterministic semantics of k-depth logics. On the one hand $B_k(\varphi|\gamma)$ counts as evidence for φ only the (syntactic representations of the) partial evaluations which suffice to obtain that φ is true, deterministically; $Pl_k(\varphi|\gamma)$, on the other hand, takes into account all the partial evaluations which do not suffice to obtain that φ is false, deterministically. Clearly, within this range of evaluations lie also all the evaluations which might make φ true, but nondeterministically.

Henceforth, for ease of visualization, given a set A, we will find it helpful at times to denote simply by $\sum A$ the expression $\sum_{a\in A} a$. With a little abuse of notation, we let now:

$$m_k(v^k_\alpha) := \sum\{m_k(\beta) \mid \beta \in Supp_\gamma(m_k),\ v^k_\beta = v^k_\alpha\}$$

and

$$I^k_\gamma := \{v^k_\alpha \mid \alpha \in Supp_\gamma(m_k)\}.$$

We obtain that

$$B_k(\varphi|\gamma) = \sum_{\alpha\in Supp_\gamma(m_k)} m_k(\alpha) v^k_\alpha(\varphi) = \sum_{v^k_\alpha \in I^k_\gamma} m_k(v^k_\alpha)\cdot v^k_\alpha(\varphi). \tag{1}$$

Hence, we can equivalently think of k-depth mass functions as actually defined over I^k_γ, which is a finer *frame of discernment* (see e.g. [27]) than $Fm_{\mathcal{L}} \cup \{*\}$, since it identifies formulas having the same k-depth consequences.

Note also that, in passing from m_0 to m_k, the frame of discernment I^k_γ gets much coarser than I^0_γ: functions which were previously distinct turn out to be the same. As an example, consider that v^0_* and $v^0_{p\vee\neg p}$ boil down to the same function, i.e. $v^1_* = v^1_{p\vee\neg p}$.

A final comment on Definition 4, and its equivalent reformulation (1), is in order. To quantify their degree of belief in a formula φ, an agent collects and "sums up" all the actual information or evidence which permits to infer φ. What our logical analysis adds to this classical view is a way of distinguishing two features of this belief formation process which are usually conflated in set-theoretic models of belief: on the one hand the representation of the evidence possessed by an agent, and on the other hand their inferential ability. As to the first feature, our model borrows the concept of *information states* from the theory of Depth-bounded Boolean logics. Information states are built starting from any formula of the language: for each α, we consider the information state where the agent holds only that α and its logical consequences (up to a fixed k) are true. This accounts for the second feature of belief formation: (limited) inferential ability. Definition 4 accounts transparently for the role of both. Confront this with the usual "possible worlds" representation which is typical of set-theoretic models of belief: each possible world corresponds to a classical evaluation, hence it requires that the truth value of each propositional variables is settled and that agents are capable of full deductive closure.

Our first proposition collects the main properties of k-depth belief functions, and indeed justifies the terminology.

Proposition 1. *Each B_k is (a-b)* normalized, *(c)* monotone *and (d)* totally monotone *function, i.e. for each formulas $\gamma, \varphi, \varphi_1, \ldots, \varphi_n$, it satisfies:*

(a) $\gamma \vdash_k \varphi$ implies $B_k(\varphi|\gamma) = 1$
(b) $\gamma \vdash_k \neg\varphi$ implies $B_k(\varphi|\gamma) = 0$
(c) $\gamma, \varphi \vdash_k \psi$ implies $B_k(\varphi|\gamma) \leq B_k(\psi|\gamma)$
(d) $B_k(\bigvee_{i=1}^n \varphi_i|\gamma) \geq \sum_{\emptyset\neq S}(-1)^{|S|-1} B_k(\bigwedge_{i\in S}\varphi_i|\gamma)$

Proof. (a). By the definition of $Supp_\gamma(m_k)$, we have $\alpha \vdash_0 \gamma$ for each $\alpha \in Supp_c(m_k)$. Hence since $\vdash_0 \subseteq \vdash_k$, we obtain $\alpha \vdash_k \gamma$, and by the transitivity of $\vdash_k$, $\alpha \vdash_k \varphi$, i.e. $v^k_\alpha(\varphi) = 1$ for each $\alpha \in Supp_\gamma(m_k)$. Finally, by the definition of B_k, we obtain

$$B_k(\varphi|\gamma) = \sum_{\alpha\in Supp_\gamma(m_k)} m_k(\alpha) v^k_\alpha(\varphi) = \sum_{\alpha\in Supp_\gamma(m_k)} m_k(\alpha) = 1.$$

(b). From the definition of B_k, we have $B_k(\varphi|\gamma) \neq 0$ if and only if there is at least a formula $\alpha \in Supp_\gamma(m_k)$, such that $\alpha \vdash_k \varphi$. On the other hand, since $\alpha \in Supp_\gamma(m_k)$, $\alpha \vdash_0 \gamma$, hence $\alpha \vdash_k \neg\gamma$. The latter together with the assumption $\gamma \vdash_k \neg\varphi$, gives us by transitivity that $\alpha \vdash_k \neg\varphi$. Hence we obtain $\alpha \vdash_k \curlywedge$, that is, $v^k_\alpha = v^k_\curlywedge$. By Definition 3 this means that $m_k(\alpha) = 0$, which is in contradiction with $\alpha \in Supp_\gamma(m_k)$.

(c) Given the definition of $B_k(\varphi|\gamma)$, it suffices to show that, whenever α is such that $v^k_\alpha(\varphi) = 1$ then $v^k_\alpha(\psi) = 1$. Assume that $v^k_\alpha(\varphi) = 1$, i.e. $\alpha \vdash_k \varphi$. Since $\alpha \in Supp_\gamma(m_k)$, we also have $\alpha \vdash_k \gamma$. On the other hand we get, by our hypothesis, $\gamma, \varphi \vdash_k \psi$, hence by the transitivity of $\vdash_k$, we obtain $\alpha \vdash_k \psi$, i.e. $v^k_\alpha(\psi) = 1$.

(d). We adapt a similar proof in Theorem 4.1 in [19]. Let us recall (see equation (1)) that m_k can be seen as a probability distribution over the set I^k_γ, and hence straightforwardly defines a probability measure (in usual, set-theoretical terms) over the finite set $\mathcal{P}(I^k_\gamma)$. We thus obtain:

$$\begin{aligned}
B_k(\bigvee_{i=1}^{n}\varphi_i|\gamma) &= \sum_{v^k_\alpha \in I^k_\gamma} m_k(v^k_\alpha) v^k_\alpha(\varphi_1 \vee \cdots \vee \varphi_n) \\
&\geq \sum_{v^k_\alpha \in I^k_\gamma} m_k(\alpha) \max(v^k_\alpha(\varphi_1), \ldots, v^k_\alpha(\varphi_n)) \\
&= m_k(\{v^k_\alpha \mid \text{for some } i,\ v^k_\alpha(\varphi_i) = 1\}) \\
&= \sum_{\emptyset \neq S \subseteq \{1,\ldots,n\}} (-1)^{|S|-1} m_k(\{v^k_\alpha \mid v^k_\alpha(\varphi_i) = 1 \text{ for all } i \in S\}) \\
&= \sum_{\emptyset \neq S \subseteq \{1,\ldots,n\}} (-1)^{|S|-1} B_k(\bigwedge_{i \in S} \varphi_i).
\end{aligned}$$

The inequality in the second line follows from the obvious fact that, for all the v^k_α such that $v^k_\alpha(\varphi) = 1$, i.e. $\alpha \vdash_k \varphi_i$, we have that $v^k_\alpha(\varphi_1 \vee \cdots \vee \varphi_n) = 1$, i.e. $\alpha \vdash_k \varphi_1 \vee \cdots \vee \varphi_n$, by the $\vee\mathcal{I}$ rule in Table 1. The third line is a reformulation of the previous one, using the additivity of m_k as a probability measure over $\mathcal{P}(I^k_\gamma)$.

The fourth line follows from the application of the inclusion-exclusion principle for probability measures. Finally, the last equality follows from the introduction and elimination rules for conjunction. Such rules determine indeed that, given any $\emptyset \neq S \subseteq \{1, \ldots, n\}$, we have that $\alpha \vdash_k \varphi_i$ for each $i \in S$, iff $\alpha \vdash_k \bigwedge_{i\in S} \varphi_i$. □

Example 2. Assume that a 0-depth bounded agent judges that the actual information it possesses provides a strong support only for the formula $p \vee q$, and none for its disjuncts. We can represent such a situation, for instance by $m_0(p \vee q) = 0.8$, $m_0(*) = 0.2$ and $m_0(\alpha) = 0$, for any other $\alpha \in Fm_\mathcal{L}$. Note in particular that $p \vee q \nvdash_0 q \vee p$ and $p \vee q \nvdash_0 (q \wedge p) \vee (q \wedge \neg p) \vee (\neg q \wedge p)$, hence one has $B_0(p \vee q) = 0.8$ and $B_0(q \vee p) = B_0((q \wedge p) \vee (q \wedge \neg p) \vee (\neg q \wedge p)) = 0$, although both $q \vee p$ and $(q \wedge p) \vee (q \wedge \neg p) \vee (\neg q \wedge p)$ classically are logically equivalent to $p \vee q$. Consider now a 1-depth agent, using the same piece of evidence, i.e. $m_1(p \vee q) = m_0(p \vee q) = 0.8$ and $m_1(*) = m_0(*) = 0.2$. For such agent $v^1_{p\vee q} = v^1_{q\vee p}$ and $v^1_{p\vee q} = v^1_{(q\wedge p)\vee(q\wedge\neg p)\vee(\neg q\wedge p)}$. Hence even though it gave no direct support for the information state on the right of both equalities, it will assign $B_1(p \vee q) = 0.8 = B_1(q \vee p)$ and $B_1(p \vee q) = 0.8 = B_1(q \wedge p) \vee (q \wedge \neg p) \vee (\neg q \wedge p)$.

4. The hierarchy of Depth-bounded Belief functions

So far, in correspondence to each k-depth bounded logic, we have defined a notion of k-depth Belief function, based on a very general definition of k-depth mass function. Whilst this responds to the natural question of grounding Belief functions on Depth-bounded Boolean logics, it still leaves important representational desiderata unaddressed. To see this, note that while each k-depth logical consequence ($k > 0$) is recursively defined in terms of logical consequences at lower depth, no analogous restriction has yet been imposed on the mass functions, as we move across different depths. In particular, given a formula α, our Definition 3 would in principle allow an agent to assign completely unrelated values to $m_0(\alpha)$ and $m_k(\alpha)$. Addressing this issue will provide a novel, to the best of our knowledge, presentation of Belief function in terms of a hierarchy of approximations thereof.

Let us recall that mass functions in Dempster-Shafer theory [27] are meant to represent an agent's judgment of evidence: in this section, we take such evidence to be determined only at a "shallow" level, corresponding to the mass function m_0. We cannot, however, just assume that each m_k equals m_0, since we need to take into account the increased logical capacity of the agent.

Indeed, m_0 is required to assign value 0 only to the α such that $v^0_\alpha = v^0_\curlywedge$. It might be the case, however, that inconsistencies are recognized as such only at a certain depth $k > 0$. Each m_k is obtained then from m_0 by distributing among noncontradictory formulas all the masses of the formulas which can be shown to be contradictory at depth k. This motivates the following definition.

Definition 5. Let $\gamma \in Fm_\mathcal{L} \cup \{*\}$ and $m_0 \colon Fm_\mathcal{L} \cup \{*\} \to [0, 1]$ be a 0-depth mass function with support $Supp_\gamma(m_k)$. We say that m_k is an *m_0-based k-depth mass function* if it satisfies the following:

- For each $\alpha \in Supp_\gamma(m_k)$, such that $v^k_\alpha \neq v^k_\curlywedge$

$$m_k(\alpha) = \frac{m_0(\alpha)}{1 - \sum\{m_0(\beta) \mid \; v^k_\beta = v^k_\curlywedge\}}.$$

Moreover, if for each $k > 0$, $Supp_\gamma(m_0) = Supp_\gamma(m_k)$, we say that the m_k are *constant m_0-based k-depth mass functions.*

The *k-depth m_0-based Belief functions* are then just obtained as in Definition 4, on the basis of m_0-based k-depth mass functions.

The m_0-based Belief functions are just special cases of those in Definition 4, hence the results of the previous section still apply.

We can think of m_0-based k-depth mass functions as arising from a peculiar kind of belief revision, which is not due to new information, but only to (higher-depth) logical reasoning. In particular, Definition 5 expresses an agent's update of degrees of belief, only upon the following new piece of information: some of the formulas to which she originally assigned non-zero mass are, in fact, contradictory. Such information could also be encoded, for instance, as a simple mass-function $m_{\curlywedge_k}$, such that

$$m_{\curlywedge_k}(\bigwedge_{\substack{\alpha \in Fm_\mathcal{L} \\ v^k_\alpha = v^k_\curlywedge}} \neg\alpha) = 1$$

and $m_{\curlywedge k}(\beta) = 0$ for any other formula. Then the m_k in Definition 5 can also be obtained by just updating m_0 with such a $m_{\curlywedge_k}$ via the Dempster-Shafer [27] rule of combination.

Note that, by Equation (1), we can also express m_k, in terms of the apparently looser constraint:

$$m_k(v^k_\alpha) = \frac{\sum\{m_0(\beta) \mid \; v^k_\beta = v^k_\alpha\}}{1 - \sum\{m_0(\beta) \mid \; v^k_\beta = v^k_\curlywedge\}} \tag{2}$$

resulting in the same class of k-depth m_0-based Belief functions.

Our next proposition shows that Belief functions which are based on constant m_0-based k-depth mass functions, determine intervals $[B_k(\varphi|\gamma), Pl_k(\varphi|\gamma)]$ for each $\varphi \in Fm_\mathcal{L}$, which get smaller as k increases. This has the welcome consequence to the effect that higher logical abilities, as measured by $\vdash_k$, lead to sharper uncertainty quantification.

Proposition 2. *Let $\gamma \in Form_\mathcal{L}$, m_0 be a 0-depth mass function, m_k be a constant m_0-based k-depth mass function with support $Supp_\gamma(m_k)$, and B_k and Pl_k the corresponding belief and plausibility functions. For each $k \geq 0$, we have $B_k(\varphi|\gamma) \leq B_{k+1}(\varphi|\gamma)$; $Pl_k(\varphi|\gamma) \geq Pl_{k+1}(\varphi|\gamma)$*

Proof. For each $\alpha \in Supp_\gamma(m_k)$, clearly $\alpha \vdash_k \varphi$ implies $\alpha \vdash_{k+1} \varphi$, hence $v^k_\alpha(\varphi) \leq v^{k+1}_\alpha(\varphi)$ for each $\varphi \in Fm_\mathcal{L}$. On the other hand, by Definition 5, it follows that for each $\alpha \in Supp_\gamma(m_k) = Supp_\gamma(m_{k+1})$ such that $v^k_\alpha \neq v^k_\curlywedge$, we have $m_k(\alpha) = m_{k+1}(\alpha)$. Hence from Definition 4, we have $B_k(\varphi|\gamma) \leq B_{k+1}(\varphi|\gamma)$. As an easy consequence of this fact, we get $Pl_k(\varphi) \geq Pl_{k+1}(\varphi)$. □

Let us consider now the connection between the k-depth Belief functions defined above and the Belief functions in the classical setting. A customary presentation of Belief functions is as set functions, see e.g. [27,13], satisfying the set-theoretic counterpart of the (a-d) in Proposition 1.

In a logical setting [19], they can be represented as functions over the boolean Lindenbaum algebra of classical logic (see e.g. [19]). Recall that elements of such algebra are the equivalence classes of the form $[\alpha]_{\equiv_\gamma}$, determined by the relation $\equiv_\gamma$ defined by $\alpha \equiv_\gamma \beta$ iff $\gamma, \alpha \vdash \beta$ and $\gamma, \beta \vdash \alpha$.

Let us now introduce, for $\alpha \in Fm_\mathcal{L}$, a function v_α, such that $v_\alpha(\varphi) = 1$ if $\alpha \vdash \varphi$ and $v_\alpha(\varphi) = 0$ otherwise. Note that distinct formulas α and β may give rise to the same valuation: we have distinct v_αs only for distinct classes of the Lindenbaum algebra $Lind_\gamma$. Hence, in analogy to what we did for Depth-bounded Belief functions, we can equivalently formulate classical Belief functions in terms of a probability distribution over the set $I_\gamma = \{v_\alpha \mid \alpha \in Fm_\mathcal{L}, \alpha \vdash \gamma\}$. More precisely, given any classical Belief function $B\colon Fm_\mathcal{L} \to [0,1]$ we can find a mass function m (also called Moebius transform) $m\colon I_\gamma \to [0,1]$ such that

$$\sum_{v_\alpha \in I_\gamma} m(v_\alpha) = 1,$$

$m(v_\curlywedge) = 0$, and the Belief function B is obtained by

$$B(\varphi|\gamma) = \sum_{v_\alpha \in I_\gamma} v_\alpha(\varphi) m(v_\alpha),$$

for any formula φ. This shows that Belief functions are obtained as convex combinations of functions (which are not classical evaluations) $v_\alpha : Fm \to \{0, 1\}$.

Against this background we can now state the main result of this section.

Theorem 2. *Let $\gamma \in Fm_{\mathcal{L}} \cup \{*\}$ and $B : Fm_{\mathcal{L}} \to [0, 1]$ be a Belief function. Then there is a sequence of Depth-bounded Belief functions B_k such that, for each φ we have*

$$B(\varphi|\gamma) = \lim_{k\to\infty} B_k(\varphi|\gamma).$$

Proof. By the argument above we have

$$B(\varphi|\gamma) = \sum_{v_\alpha \in I_\gamma} m(v_\alpha) v_\alpha(\varphi),$$

so it suffices to take any mass function m_0 such that, for each $v_\alpha \in I_\gamma$

$$\sum \left\{ m_0(\beta) \mid \beta \in Fm_{\mathcal{L}}, v_\alpha = v_\beta \right\} = m(v_\alpha).$$

Clearly, since $\sum_{v_\alpha \in I_\gamma} m(v_\alpha) = 1$ we also get

$$\sum_{\alpha \in Supp_\gamma(m_0)} m_0(\alpha) = 1.$$

Moreover, by the definition of m_0, since $m(v^0_{\curlywedge}) = 0$, we get $m_0(\alpha) = 0$ for each $v_\alpha = v_{\curlywedge}$. Hence m_0 is a 0-depth mass function and, letting m_k be the corresponding m_0-based k-depth mass function (see Definition 5 and the reformulation thereafter in Equation (2)), we have

$$m_k(v^k_\alpha) = \sum \left\{ m_0(\beta) \mid \; v^k_\beta = v^k_\alpha \right\}.$$

By Theorem 1 we know that $\lim_{k\to\infty} v^k_\alpha(\varphi) = v_\alpha(\varphi)$, hence we get

$$\begin{aligned} \lim_{k\to\infty} m_k(v^k_\alpha) &= \lim_{k\to\infty} \sum \{ m_0(\beta) \mid \; v^k_\beta = v^k_\alpha \} \\ &= \sum \{ m_0(\beta) \mid \; v_\beta = v_\alpha \} \\ &= m(v_\alpha). \end{aligned}$$

We finally obtain

$$\begin{aligned} \lim_{k\to\infty} B_k(\varphi|\gamma) &= \lim_{k\to\infty} \sum_{\alpha \in Supp_\gamma(m_k)} v^k_\alpha(\varphi) m_k(v^k_\alpha) \\ &= \sum_{v_\alpha \in I_\gamma} v_\alpha(\varphi) \cdot m(v_\alpha) = B(\varphi|\gamma). \quad \square \end{aligned}$$

To sum up, the conditions on the mass function pinned down by Definition 5 lead to the construction of a hierarchy of Belief functions which approximate Dempster-Shafer Belief functions. Each element in the hierarchy identified by Theorem 2 can thus be interpreted as an approximation of the Dempster-Shafer degree of belief of a *realistic* agent, i.e. one whose logical abilities are bounded by $\vdash_k$.

5. Restricting the focal formulas: atoms and subatoms

Let us now investigate some classes of k-depth Belief functions arising from restrictions of the set of focal formulas. Throughout this section, we fix $Var_{\mathcal{L}} = \{p_1, \ldots, p_n\}$ and assume, for simplicity, no background information, or, which is the same, that the background information amounts to $*$.

We first consider the case of $Supp(m_0) \subseteq At_{\mathcal{L}}$, where

$$At_{\mathcal{L}} = \{\bigwedge \pm p_i \mid p_i \in Var_{\mathcal{L}}\}.$$

In the classical setting, this reduces Belief functions to probabilities. Atoms correspond indeed to boolean evaluations: it is easy to see that for each boolean evaluation $v : Fm \to \{0, 1\}$ there is a unique atom α such that for all $\varphi \in Fm$, we have $\alpha \vdash \varphi$ iff $v(\varphi) = 1$, i.e. $v(\varphi) = v_\alpha(\varphi)$.

The restriction to atoms produces also a collapse of the hierarchy of the Depth-bounded logics: atoms correspond indeed to evaluations where (informational) truth and falsity is assigned to all propositional variables, and the tables in Fig. 1 coincide with the classical ones, when the indeterminate value $*$ is not assigned. As a proof-theoretical counterpart of this fact, it is easy to check that, for any atom $\alpha \in At_{\mathcal{L}}$, for each $k \geq 0$, we have $\alpha \vdash_k \varphi$ iff $\alpha \vdash \varphi$ for every $\varphi \in Fm_{\mathcal{L}}$, i.e. $v^k_\alpha = v_\alpha$.

In light of such considerations, we obtain that for any 0-depth mass assignment m_0, with $Supp(m_0) \subseteq At_{\mathcal{L}}$, the corresponding m_0-based k-depth Belief function is such that $B_k(\varphi) = B_{k+1}(\varphi)$ for each $k \geq 0$ and for each $\varphi \in Fm_{\mathcal{L}}$. Each B_k thus satisfies, in addition to the properties of Lemma 1, also finite additivity, and therefore is just as a probability function.

A more interesting example arises, when letting

$$Subat_{\mathcal{L}} = \{\bigwedge_{i \in I} \pm p_i \mid I \subseteq \{1, \dots, n\}\} \cup \{*\}$$

we assume that $Supp(m_0) \subseteq Subat_{\mathcal{L}}$. While atoms are in one-to-one correspondence with boolean evaluations, each subatom α determines a corresponding three-valued non-deterministic evaluation, i.e. the one which assigns the value $*$ to each propositional variable not occurring in α.

Given this correspondence, we can take $m_0(\alpha)$ to represent the belief that an agent assigns exclusively to a single (three-valued) evaluation, i.e. the one corresponding to the formula α. This amounts to the agent being in possess of information about the truth or falsity of some propositional variables, but not necessarily all of them. This case is intermediate between the general setting presented in Section 6, where the agent expresses belief over arbitrary formulas (which we can think of as belief committed to sets of partial evaluations) and the case just discussed above, reducing to probability, where the agent expresses belief over a classical evaluation, i.e. a belief concerning the truth or falsity of every propositional variable in the language.

Let us observe a few interesting features of this family of k-depth Belief functions.

First, unlike the previous case based on atoms, the k-depth Belief functions B_k will be in general distinct from each other.

Second, the B_0 Belief functions turn out to be finitely additive, as we show in the following.

Proposition 3. *Let m_0 be a 0-depth mass function, such that $Supp_\gamma(m_0) \subseteq Subat_{\mathcal{L}}$. For each formulas $\varphi, \psi \in Fm_{\mathcal{L}}$, we have*

$$B_0(\varphi \vee \psi) = B_0(\varphi) + B_0(\psi) - B_0(\varphi \wedge \psi).$$

Proof. By definition

$$B_0(\varphi \vee \psi) = \sum_{\alpha \in Supp(m_0)} m_0(a) v^0_\alpha(\varphi \vee \psi) = \sum_{\alpha \in Subat_{\mathcal{L}}} m_0(a) v^0_\alpha(\varphi \vee \psi).$$

Note that $\vdash_0$ has the disjunction property, i.e. $\varphi \vee \psi$ is derivable iff either φ or ψ are derivable. This means that, if $\alpha \in Subat_{\mathcal{L}}$, then $v^0_\alpha(\varphi \vee \psi) = 1$ iff $v^0_\alpha(\varphi) = 1$ or $v^0_\alpha(\psi) = 1$. We obtain then

$$v^0_\alpha(\varphi \vee \psi) = v^0_\alpha(\varphi)(1 - v^0_\alpha(\psi)) + v^0_\alpha(\psi) \cdot (1 - v^0_\alpha(\varphi)) + v^0_\alpha(\varphi) \cdot v^0_\alpha(\psi).$$

Hence, using the latter, we get:

$$\begin{aligned} B_0(\varphi \vee \psi) &= \sum_{\alpha \in Subat_{\mathcal{L}}} m_0(v^0_\alpha) v^0_\alpha(\varphi \vee \psi) \\ &= \sum_{\alpha \in Subat_{\mathcal{L}}} m_0(\alpha)[v^0_\alpha(\varphi) + v^0_\alpha(\psi) - v^0_\alpha(\varphi) \cdot v^0_\alpha(\psi)] \\ &= \sum_{\alpha \in Subat_{\mathcal{L}}} m_0(\alpha) v^0_\alpha(\varphi) + \sum_{\alpha \in Subat_{\mathcal{L}}} m_0(\alpha) v^0_\alpha(\psi) \\ &- \sum_{\alpha \in Subat_{\mathcal{L}}} m_0(\alpha) v^0_\alpha(\varphi \wedge \psi) \\ &= B_0(\varphi) + B_0(\psi) - B_0(\varphi \wedge \psi). \quad \square \end{aligned}$$

Remark 1. Note that the result for B_0 is also an immediate consequence of Theorem 5 in [20], since the evaluations v^0_α satisfy T2 and T3, in the terminology of [20].

Let us see now some other examples of m_0-based k-depth Belief functions based on the support $Subat_{\mathcal{L}}$.

Example 3. Let $Var_{\mathcal{L}} = \{p, q\}$ and m_0 be a 0-depth mass function, which is uniformly distributed over the set

$$Supp(m_0) = Subat_{\mathcal{L}} = \{*, p, \neg p, q, \neg q, p \wedge q, p \wedge \neg q, \neg p \wedge q, \neg p \wedge \neg q\}$$

i.e. we let $m_0(a) = 1/9$ for each formula $\alpha \in Supp(m_0)$. It is easy to see that, if α is one of $p, q, p \wedge q, p \wedge \neg q, \neg p \wedge q$ then $\alpha \vdash_0 p \vee q$. On the other hand, only if α is $\neg p \wedge \neg q$ then $\alpha \vdash_0 \neg(p \vee q)$. Hence by the definition of B_0, we have:

$$B_0(p \vee q) = 5/9 \quad B_0(\neg(p \vee q)) = 1/9 \quad Pl_0(p \vee q) = 8/9.$$

Recall that, under a uniform probability distribution over atoms in the classical case, we would get $P(p \vee q) = 3/4 \in [5/9, 8/9]$.

Note that $\alpha \vdash_k p \vee q$ and $\alpha \vdash_k \neg(p \vee q)$ holds exactly in the same cases as for $\vdash_0$, hence we have

$$B_k(p \vee q) = B_0((p \vee q) = 5/9 \quad B_k(\neg(p \vee q)) = B_0(\neg(p \vee q) = 1/9$$
$$Pl_k(p \vee q) = Pl_0(p \vee q) = 8/9,$$

for each $k \geq 1$.

Example 4. Let us consider the formula $p \vee \neg p$, with a mass function m_0 such that $Supp(m_0) = \{*, p, \neg p\}$. We have:

$$p \vdash_0 p \vee \neg p \quad \neg p \vdash_0 p \vee \neg p \qquad \emptyset \nvdash_0 p \vee \neg p$$
$$p \vdash_1 p \vee \neg p \quad \neg p \vdash_1 p \vee \neg p \qquad \emptyset \vdash_1 p \vee \neg p$$

We obtain:

$$B_0(p \vee \neg p) = m_0(p) + m_0(\neg p) \quad B_0(\neg(p \vee \neg p)) = m_0(v_{\curlywedge}) = 0$$
$$Pl_0(p \vee \neg p) = m_0(p) + m_0(\neg p) + m_0(*) = 1$$

On the other hand:

$$B_1(p \vee \neg p) = m_0(p) + m_0(\neg p) + m_0(*) = 1$$
$$B_1(\neg(p \vee \neg p)) = 0 \quad Pl_1(p \vee \neg p) = 1$$

We have thus $B_0(p \vee \neg p) = B_0(p) + B_0(\neg p)$, while $B_1(p \vee \neg p) > B_1(p) + B_1(\neg p) - B_1(\neg(p \vee \neg p)$.

The last example helps to illustrate the reason why B_0 behaves as a finitely additive measure, while the B_k with $k > 0$ are superadditive. As we have seen, $B_1(p \vee \neg p)$ includes the belief in $m(*)$, which is only "virtually" in favour of $p \vee \neg p$: neither $B_1(p)$ nor $B_1(\neg p)$ include such belief. The example also shows that k-depth Belief functions are a proper subclass of Belief functions, due to the lack of normalization on classical tautologies: one can have indeed $B_0(p \vee \neg p) < 1$ while this cannot occur for classical Belief functions.

6. Approximating classical probability

In this section we suggest yet another way to restrict the focal formulas of k-depth mass functions, resulting in sequences of Depth-bounded Belief functions which approximate classical probability functions.

In Section 4 and 5, we assumed that the set of focal formulas $Supp_{\gamma}(m_k)$ for some $\gamma \in Fm_{\mathcal{L}} \cup \{*\}$ and $k > 0$ was the same as $Supp_{\gamma}(m_0)$, except for the possible removal (and redistribution of the mass) of formulas which turn out to be inconsistent at depth k. We will now provide some more restrictive constraints on $Supp_{\gamma}(m_k)$.

The guiding idea is that $Supp_{\gamma}(m_0)$ should only contain formulas which an agent is able to judge already at a shallow level. As the depth increases, we assume that, besides higher inferential capacity, the agents have also higher "imaginative" capacity, i.e. capacity for weighting the uncertainty of scenarios not immediately given to them (see [21] for a related approach).

Let us now give our formal definition of the k-depth mass function, with focal formulas defined according to the ideas sketched above.

Definition 6 *(k-depth revising mass function).* Let $\gamma \in Fm_{\mathcal{L}} \cup \{*\}$ and m_0 be a 0-depth mass function with support $Supp_{\gamma}(m_0)$. We say that the functions m_k for $k > 0$ are *m_0-based k-depth revising mass functions* iff

1. $Supp_{\gamma}(m_{k+1})$, for $k \geq 0$, contains only formulas of the kind $\alpha \wedge p_i$ and $\alpha \wedge \neg p_i$, for each $\alpha \in Supp_{\gamma}(m_k)$, $p_i \in Var_{\mathcal{L}}$, provided that $\pm p_i$ does not occur already as one of the conjuncts of α.

2. For each $\alpha \in Supp_\gamma(m_k)$

$$\sum \left\{ m_{k+1}(\alpha \wedge \pm p_i) \mid \alpha \wedge \pm p_i \in Supp_\gamma(m_{k+1}) \right\} = m_k(\alpha).$$

3. For each $\alpha \in Supp_\gamma(m_k)$,

$$m_{k+1}(\alpha \wedge \pm p_i) = 0 \text{ if } \alpha \wedge \pm p_i \vdash_{k+1} \curlywedge.$$

Note that the second condition in the definition above amounts to redistributing, for each $\alpha \in Supp_\gamma(m_k)$ the mass $m_k(\alpha)$ among the formulas $\alpha \wedge \pm p_i \in Supp_\gamma(m_{k+1})$. It is easy to check that

$$\sum_{\alpha \wedge \pm p_i \in Supp_\gamma(m_{k+1})} m_{k+1}(\alpha \wedge \pm p_i) = 1,$$

hence the m_0-based k-depth revising mass functions are just a particular case of k-depth mass functions, which determine corresponding Belief functions $B_k(|\gamma)$ by Definition 4. All the results in Section 3 will still hold.

Example 5. Consider a 0-depth mass function such that $Supp(m_0) = \{p \vee q, *\}$ $Var_\mathcal{L} = \{p, q\}$. Let us determine the support of the corresponding m_0-based k-depth revising mass functions m_k and Belief functions B_k. We have

$$Supp(m_1) = \{(p \vee q) \wedge p, (p \vee q) \wedge \neg p, (p \vee q) \wedge q, (p \vee q) \wedge \neg q\} \cup \{p, \neg p, q, \neg q\}$$

At depth 2, we get

$$\begin{aligned} Supp(m_2) = \{&p \wedge q, p \wedge \neg q, \neg p \wedge q, \neg p \wedge \neg q, q \wedge p, q \wedge \neg p, \\ &q \wedge p, q \wedge \neg q\} \cup \{(p \vee q) \wedge p \wedge q, (p \vee q) \wedge \neg p \wedge q, \\ &(p \vee q) \wedge q \wedge p, (p \vee q) \wedge q \wedge \neg p, (p \vee q) \wedge \neg q \wedge p\}. \end{aligned}$$

All the formulas in $Supp(m_2)$ are actually already 2-depth logically equivalent to atoms (with many repetitions of logically equivalent formulas), hence B_2 is just a probability function.

Our last result shows that Definition 6 identifies the conditions that mass functions must satisfy in order for k-depth Belief functions to approximate classical probability functions.

Theorem 3. *Let $P : Fm_\mathcal{L} \to [0, 1]$ be a classical probability function and $\gamma \in Fm_\mathcal{L} \cup \{*\}$. Then there is a sequence of revising Depth-bounded Belief functions B_k such that, for each $\varphi \in Fm_\mathcal{L}$, we have $P(\varphi|\gamma) = \lim_{k\to\infty} B_k(\varphi|\gamma)$.*

Proof. First, we know that

$$P(\varphi|\gamma) = \sum_{v_\alpha \in I_\gamma} m(v_\alpha) v_\alpha(\varphi)$$

where $I_\gamma \subseteq \{v_\alpha \mid \alpha \in At_\mathcal{L}\}$ and $m(v_\alpha) = P(\alpha|\gamma)$. Let us take the mass function m_0 such that $Supp_\gamma(m_0) = \{\gamma\}$, and $m_0(\gamma) = 1$. Since our language is finite and $|Var_\mathcal{L}| = n$, we will obtain by construction that, for any m_0-based n-depth revising mass function, all the formulas in $Supp_\gamma(m_n)$ are n-depth logically equivalent to atoms. Let now m_n be any such mass function, that satisfies:

$$\sum \left\{ m_n(\beta) \mid \beta \in Supp_\gamma(m_n)\ ,\ v^n_\beta = v^n_\alpha \right\} = m(v_\alpha)$$

for each $\alpha \in At_\mathcal{L}$. All the remaining mass functions m_i for $i < n$ are then uniquely determined, by Definition 6. Now, for all $B_k(\varphi|\gamma)$ with $k \geq n$, we get $B_k(\varphi|\gamma) = B_n(\varphi|\gamma)$, hence:

$$lim_{k\to\infty} B_k(\varphi|\gamma) = B_n(\varphi|\gamma) = P(\varphi|\gamma). \quad \square$$

This result sheds new light on the relation between probability functions and Belief functions. It has long been known that probability functions are a special case of Belief functions in which the unit mass is distributed on all the “singletons” of a suitable sample space. Our non-classical logical setting uncovers a further dimension of comparison which depends on the rather subtle interplay between the information possessed by an agent and their inferential abilities.

7. Future work

As recalled in Section 2 it is shown in [4] and [3] that the informational semantics of Depth-bounded Boolean logics forms the basis to define an infinite hierarchy of *tractable* deductive systems (with no syntactic restriction on the language adopted) whose upper limit coincides with classical propositional logic. Hence this hierarchy can be seen as a tractable approximation to classical logic. It is therefore natural to ask whether the hierarchy of Depth-bounded Belief functions inherit this highly desirable feature.

Unfortunately, grounding Belief functions on Depth-bounded consequence relations is not outright sufficient to make the representation of belief itself feasible. Indeed the opposite is true. Our description of how a bounded agent would quantify their belief is likely to be computationally harder than the description of the ideal case, based on classical logic. Our k-depth mass functions require indeed to specify a much larger number of values than required by the classical case. Note that this phenomenon occurs in a sense already where the underlying logic is classical. While we can arguably consider Belief functions to be less idealized models of uncertainty than probabilities, to the extent that they treat knowledge and ignorance asymmetrically, defining the latter functions on a language with n propositional variables requires to specify $2^n - 1$ values, while for the former one needs $2^{2^n} - 1$ (see [19]). Already in the classical setting, computational complexity is recognized as one of the soft spots of Belief functions, since in particular combining Belief functions is known to be $\#P$-complete [18]. Still unfeasible, though computationally better, is the problem of deciding whether there is a Belief function satisfying some constraints, given in the form of a system of linear polynomials. This is indeed shown to be NP-complete, see [19].

In practice various methods exist for reducing the computational complexity for application purposes, among them the use of Monte Carlo methods: an overview of efficient algorithms can be found in [29]. In particular, some works [1,2] show how a dramatic reduction of complexity bounds (from exponential to linear) results from imposing constraints on the focal sets. This gives a good hint that the Belief functions defined in Section 5 and 6 are a promising starting point to obtain a more tractable treatment of uncertainty, since they add to the polynomial-time decidability of the underlying consequence relation $\vdash_k$, a stricter control over the set of focal formulas. But the full-fledged investigation on how to exploit the hierarchy of approximations of Belief functions to get a computationally feasible quantification of uncertainty must be postponed to future work.

Future work will also relate the hierarchy of Depth-bounded Belief functions to the general coherence-based approach of [11], therefore providing tighter connections with the very idea of rational quantification of uncertainty. As a promising starting point, note that each $B_k(\alpha|\gamma)$ in Definition 4 is a convex combination of evaluations. It follows from results in [20] that $B_k(\varphi|\gamma)$ does not permit a Dutch Book, i.e. there do not exist formulas $\varphi_1, \dots, \varphi_n$ in $Fm_{\mathcal{L}}$ and odds $s_1, \dots, s_n \in \mathbb{R}$ such that for all $v_\alpha^k \in I_\gamma^k$

$$\sum_{i=1}^{n} s_i(v_\alpha^k(\varphi_i) - B_k(\varphi_i|\gamma)) < 0.$$

This clearly paves the way for a detailed analysis of the coherence conditions of each element in the hierarchy of Depth-bounded Belief functions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J.A. Barnett, Computational methods for a mathematical theory of evidence, in: Classic Works of the Dempster-Shafer Theory of Belief Functions, in: Studies in Fuzziness and Soft Computing, vol. 219, Springer, 2008, pp. 197–216.
[2] W.T. da Silva, R.L. Milidiú, Algorithms for combining belief functions, Int. J. Approx. Reason. 7 (1–2) (1992) 73–94.
[3] M. D'Agostino, An informational view of classical logic, Theor. Comput. Sci. 606 (2015) 79–97.
[4] M. D'Agostino, M. Finger, D.M. Gabbay, Semantics and proof-theory of depth-bounded Boolean logics, Theor. Comput. Sci. 480 (2013) 43–68.
[5] M. D'Agostino, T. Flaminio, H. Hosni, Rational beliefs real agents can have - a logical point of view, in: IDM@NIPS, in: Proceedings of Machine Learning Research, vol. 58, PMLR, 2016, pp. 97–109.
[6] M. D'Agostino, L. Floridi, The enduring scandal of deduction, Synthese 167 (2) (2009) 271–315.
[7] B. de Finetti, Sul significato soggettivo della probabilità, Fundam. Math. 17 (1931) 289–329.
[8] T. Denoeux, 40 years of dempster-Shafer theory, Int. J. Approx. Reason. 79 (2016) 1–6.
[9] D. Ellsberg, Risk, ambiguity, and the Savage axioms, Q. J. Econ. 75 (4) (1961) 643–669.
[10] T. Flaminio, L. Godo, H. Hosni, On the logical structure of de Finetti's notion of event, J. Appl. Log. 12 (3) (2014) 279–301.
[11] T. Flaminio, L. Godo, H. Hosni, Coherence in the aggregate: a betting method for belief functions on many-valued events, Int. J. Approx. Reason. 58 (2015) 71–86.
[12] I. Gilboa, Theory of Decision Under Uncertainty, Cambridge University Press, 2009.
[13] J.Y. Halpern, Reasoning About Uncertainty, MIT Press, 2005.
[14] J.M. Keynes, A Treatise on Probability, Harper & Row, 1921.
[15] F.H. Knight, Risk, Uncertainty and Profit, Beard Books Inc., 1921.

[16] D.V. Lindley, Understanding Uncertainty, John Wiley and Sons, 2006.
[17] D. Mundici, Advanced Lukasiewicz Calculus and MV-Algebras, Springer, 2011.
[18] P. Orponen, Dempster's rule of combination is #p-complete, Artif. Intell. 44 (1–2) (1990) 245–253.
[19] J.B. Paris, The Uncertain Reasoner's Companion: A Mathematical Perspective, Cambridge University Press, 1994.
[20] J.B. Paris, A note on the Dutch book method, in: Gert De Cooman, Terrence Fine, Teddy Seidenfeld (Eds.), ISIPTA'01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications, Ithaca, NY, USA, Shaker, 2001, pp. 301–306.
[21] J.B. Paris, A. Vencovská, A model of belief, Artif. Intell. 64 (2) (1993) 197–241.
[22] G. Parmigiani, L. Inoue, Decision Theory: Principles and Approaches, Wiley, 2009.
[23] J. Pearl, Probabilistic Reasoning in Intelligent Systems - Networks of Plausible Inference, Morgan Kaufmann Series in Representation and Reasoning, Morgan Kaufmann, 1989.
[24] W.V. Quine, The roots of reference, Br. J. Philos. Sci. 27 (1) (1976) 93–96.
[25] L.J. Savage, The Foundations of Statistics, 2nd edition, Dover, 1972.
[26] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (3) (1989) 571–587.
[27] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[28] G. Shafer, Perspectives on the theory and practice of belief functions, Int. J. Approx. Reason. 4 (5–6) (1990) 323–362.
[29] N. Wilson, Algorithms for Dempster-Shafer Theory, Springer, Netherlands, Dordrecht, 2000, pp. 421–475.