Logic-Based Approximations of Preferences

Paolo Baldi(✉)

Department of Human Studies, University of Salento, Lecce, Italy
paolo.baldi@unisalento.it

Abstract. In this exploratory work, we provide a general framework, based on Depth-Bounded Boolean logic, for addressing some of the criticisms towards Savage's approach to the foundations of decision theory. We introduce a sequence of approximating preferences structures and show that, under suitable conditions such preferences give rise to a qualitative probability which is almost representable by a finitely additive probability.

1 Introduction

In his seminal work, first published in 1954, and revisited in 1972 [15], Savage laid down a foundational framework for decision-making under uncertainty. His system is based on acts, which are rendered as functions mapping states into outcomes, and on preferences on such acts, which need to obey certain rationality axioms.

Savage's general setup, as well as his axioms, have been since subjected to wide scrutiny and criticisms. Much controversy has been raised in particular on the so-called *Sure-Thing Principle* (STP), that allows an agent to reach a preference by decomposing it in preferences over two mutually exclusive and jointly exhaustive subcases. In Savage's words, the principle is motivated as follows:

> A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew that the Democratic candidate were going to win, and decides that he would. Similarly, he considers whether he would buy if he knew that the Republican candidate were going to win, and again finds that he would. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say [15].

The purpose of this work is to provide a logical perspective, both on Savage's well-known framework [15] for the foundation of decision theory, and on its criticisms, arising from the famous scenarios presented by Ellsberg [10] and

V. Malvone and A. Murano (Eds.): EUMAS 2023, LNAI 14282, pp. 280–294, 2023.
https://doi.org/10.1007/978-3-031-43264-4_18

Allais [1]. Both of these scenarios provided patterns of preferences deemed plausible, and yet conflicting with Savage's axioms, in particular with the Sure-Thing Principle.

The key observation behind this work is the similarity of STP with what in classical logic is known as the *Principle of Bivalence* (PB). To clarify the meaning of PB, we first present it as a rule in natural-deduction style, as follows [9]:

$$\dfrac{\begin{matrix}[\varphi] \\ \vdots \\ \psi\end{matrix} \qquad \begin{matrix}[\neg\varphi] \\ \vdots \\ \psi\end{matrix}}{\psi}\ (\text{PB})$$

meaning that, to infer the formula ψ, it suffices to infer it both under the assumption that φ is the case and under the assumption that $\neg\varphi$ is the case. The square brackets around the formulas φ and $\neg\varphi$ signal that those are pieces of information assumed for the sake of deriving ψ, but not actually held true (they are *discharged*, in natural deduction terminology). Following [4], we call this type of information *hypothetical*, in contrast to the *actual* information held by an agent. Let us note that the inference rule (PB) is also called a "logical" sure-thing principle in [2], where analogies and differences with STP are analyzed. In particular, [2] stresses that "STP is a desideratum of rational behavior, but not logically necessary", as is the case instead for PB.

In the light of the development of various non-classical logics, considering PB as logically necessary, without further qualification, is not enough. In particular, choosing suitable pieces of hypothetical information for its application in logical deductions, is a complex matter. This may play an important role in decision-making, as we illustrate in the following.

Example 1. You have an urn with balls that are numbered 1–100, and are colored in unknown proportions. Three balls with numbers x_1, x_2, x_3 are extracted from the urn. You are told that $x_2 = x_1 + 1$ and $x_3 = x_2 + 1$. Ball number x_1 is red and ball number x_3 is blue. You have to choose among the following:

– h: You earn 100 euro, if $x_2 = x_1 + 1$ and $x_3 = x_2 + 1$, 0 otherwise.
– h': You earn 110 euro if it holds that, among the extracted balls
  (δ): "a red ball and a non-red ball have numbers that differ by 1", 0 otherwise.

The information provided is sufficient to assess that h always returns the payoff 100. It might be however less obvious that also h' will always return the highest payoff 110. It suffices to reason by cases: if x_2 is red, then, since $x_3 = x_2 + 1$ and x_3 is not red, δ holds. On the other hand, if x_2 is not red, since $x_2 = x_1 + 1$ and x_1 is red, δ still holds.

We find it plausible that agents might prefer h to h', although the payoff for h' is higher than that for h, and both are certain for the agent. In support of this conjecture, note that in empirical research [18], under similar information, over 80% of subjects claimed that it is impossible to determine whether an assertion of the same logical form as δ is true.

We might say that, in the above example, an agent preferring h to h' is behaving irrationally, or is perhaps attributing a *cost* to the very act of doing inferences, a cost which is not immediately captured neither by classical logic, nor by Savage's standard decision-theoretic framework.

PB is indeed costly for realistic agents, and bounding its use makes logical inference tractable, in the sense of computational complexity [17], in contrast to the intractability (under the usual $P \neq NP$ assumption) of classical propositional logic.

This observation is at the core of a family of logical systems, dubbed Depth-Bounded Boolean logics [8] (DBBLs), which allow only for a limited application of PB, and provide tractable approximations of classical logic.

Building on previous work on uncertainty measures in DBBLs [3,4], we introduce in the following a sequence of preferences approximating Savage's framework, which are based on the limited use of PB and hypothetical information.

This setting allows us to provide a unified account of Savage's axioms, and of the preferences in Allais, in Ellsberg, and in Example 1 above. All such preferences will be considered indeed compatible with (our reformulation of) Savage's axioms, and in particular with the Sure-thing principle, but only at the lowest level of our sequence, where no use of hypothetical information is permitted. Furthermore, following Savage, we show that the sequence of approximating preferences determines a finitely additive measure, in the limit.

The paper is further structured as follows. In Sect. 2 we present our analysis of actual and hypothetical information, based on DBBLs. Section 3 introduces our sequence of approximating preference relations, provides a reformulation of some of Savage's basic axioms in that setting, and analyzes our main examples. Section 4 provides the conditions under which the sequence of approximating preferences determines a finitely additive measure in the limit. Finally, we provide some conclusions and hints at future work.

2 Hypothetical and Actual Information

Before proceeding, we briefly fix some notation. We consider a propositional logical language $\mathcal{L}$, with the usual classical connectives $\wedge, \vee, \rightarrow, \neg$ and set of propositional variables $\{p_1, \dots, p_n, \dots\}$. The set of formulas will be denoted by Fm, and lowercase Greek letters will be used to refer to formulas. We denote by $S(\varphi)$ the set of subformulas of φ.

We now recall some crucial ideas of the DBBLs, mentioned in the introduction. These logics permit to distinguish between actual and hypothetical information in logical deduction, and determine a hierarchy, with a parameter k measuring the amount of allowed nested use of hypothetical information.

The 0-depth logic, which will be our main focus here, is a logic that does not allow any application of PB, and is thus concerned only with the manipulation of *actual* information. This logic is proof-theoretically defined in terms of the INTroduction and ELIMination (INTELIM) rules in Table 1. The rules are defined for each connective, both when occurring positively (as the main connective of a formula) and negatively (in the scope of a negation).

Table 1. Introduction and Elimination rules.

$$\frac{\varphi \quad \psi}{\varphi \wedge \psi}\ (\wedge\mathcal{I}) \qquad \frac{\neg\varphi}{\neg(\varphi \wedge \psi)}\ (\neg\wedge\mathcal{I}1)$$

$$\frac{\neg\psi}{\neg(\varphi \wedge \psi)}\ (\neg\wedge\mathcal{I}2) \qquad \frac{\neg\varphi \quad \neg\psi}{\neg(\varphi \vee \psi)}\ (\neg\vee\mathcal{I})$$

$$\frac{\varphi}{\varphi \vee \psi}\ (\vee\mathcal{I}1) \qquad \frac{\psi}{\varphi \vee \psi}\ (\vee\mathcal{I}2)$$

$$\frac{\varphi \quad \neg\varphi}{\bot}\ (\bot\mathcal{I}) \qquad \frac{\varphi}{\neg\neg\varphi}\ (\neg\neg\mathcal{I})$$

$$\frac{\varphi \vee \psi \quad \neg\varphi}{\psi}\ (\vee\mathcal{E}1) \qquad \frac{\varphi \vee \psi \quad \neg\psi}{\varphi}\ (\vee\mathcal{E}2)$$

$$\frac{\neg(\varphi \vee \psi)}{\neg\varphi}\ (\neg\vee\mathcal{E}1) \qquad \frac{\neg(\varphi \vee \psi)}{\neg\psi}\ (\neg\vee\mathcal{E}2)$$

$$\frac{\varphi \wedge \psi}{\varphi}\ (\wedge\mathcal{E}1) \qquad \frac{\varphi \wedge \psi}{\psi}\ (\wedge\mathcal{E}2)$$

$$\frac{\neg(\varphi \wedge \psi) \quad \varphi}{\neg\psi}\ (\neg\wedge\mathcal{E}1) \qquad \frac{\neg(\varphi \wedge \psi) \quad \psi}{\neg\varphi}\ (\neg\wedge\mathcal{E}2)$$

$$\frac{\neg\neg\varphi}{\varphi}\ (\neg\neg\mathcal{E}) \qquad \frac{\bot}{\varphi}\ (\bot\mathcal{E})$$

We note in passing that the logic has also a non-deterministic semantics, with evaluations capturing the *information* actually held by an agent rather than *truth*, as a primitive notion [7].

The rules encode the principles for the manipulation of information actually possessed by an agent, for each of the connectives of the language. We refer to [7,8] for further details and motivation. The 0-depth consequence relation is defined as follows.

Definition 1. *Let $T \cup \{\varphi\} \subseteq Fm$. $T \vdash_0 \varphi$ if there is a sequence of formulas $\varphi_1, \ldots, \varphi_m$ such that $\varphi_m = \varphi$ and each φ_i is either in T or it is obtained by an application of the rules in Table 1 from formulas φ_j with $j < i$.*

Note that, by direct inspection of the rules in Table 1, we have $\nvdash_0 p \vee \neg p$. In fact, this logic, which is strictly weaker than classical logic, has no tautologies at all. The relation $\vdash_0$ captures inferences that are "trivial" in their reliance solely on actual information. This is also reflected computationally, by the fact that, in contrast to classical propositional logic, $\vdash_0$ can be checked in polynomial time [8].

While 0-depth logic permits only to represent actual information, and lack thereof, classical logical proofs also involve reasoning about hypothetical information. Consider again $\nvdash_0 p \vee \neg p$. It can be easily shown that, on the other hand, $p \vdash_0 p \vee \neg p$ and $\neg p \vdash_0 p \vee \neg p$. Hence, we can show that $p \vee \neg p$ is derivable just by one application of PB, using the hypothetical information p and $\neg p$. In DBBLs this amounts to saying that $\vdash_1 p \vee \neg p$. The consequence $\vdash_k$ for $k > 0$ is formally defined as follows, see also [8].

Definition 2. *Let $k > 0$. Then $T \vdash_k \varphi$ if there is a $\psi \in S(T \cup \{\varphi\})$ such that $T, \psi \vdash_{k-1} \varphi$ and $T, \neg\psi \vdash_{k-1} \varphi$.*

The parameter k is thus a "counter" which keeps track of how many nested instances of reasoning by cases are needed for the agent to decide a sentence of interest.

In this work we use only 0-depth logics, to deal with actual information, alongside with a sequence of (depth-bounded) forests, to represent the further hypothetical information which may be used by an agent.

Let us recall the notion of depth-bounded forests, in a slightly modified form from [4]. We start with a set $Supp \subseteq Fm \cup \{*\}$, which represents the information explicitly provided to the agent. The symbol $*$ is meant to the represent the absence of any information. *Supp* collects background information, which may be of the form "γ holds", or "the probability of γ_i is p_i" where p_i may be the frequency or objective chance of γ_i. If no such information is available to the agent, we let $Supp = \{*\}$. We further impose that for any $\alpha, \beta \in Supp$, such that $\alpha \neq *, \beta \neq *$ we have[1] $\alpha, \beta \vdash_0 \bot$.

Depth-bounded forests are built, starting from *Supp* and suitably expanding the nodes with two new children nodes, representing an instance of PB obtained by considering a certain piece of hypothetical information and its negation, respectively.

In the following, for any formula $\gamma \in Fm$, we say that γ *decides* δ if $\gamma \vdash_0 \delta$ or $\gamma \vdash_0 \neg\delta$. By the *depth of a node* in a forest, in the usual graph-theoretic sense, we mean the length of the path from the root of a tree in the forest to the node. We then say that a leaf α is *closed* if α 0-decides each formula $\delta \in S(\alpha)$. A leaf which is not closed is said to be *open*.

Definition 3. *Let $Supp \subseteq Fm \cup \{*\}$. We define recursively,* a sequence $(F_k)_{k \in \mathbb{N}}$ of depth-bounded forests based on *Supp (DBF, for short), as follows:*

1. *For $k = 0$ we let F_0 be a forest with no edges, and with the set of vertices equal to Supp*[2].
2. *The forest F_k, for $k \geq 1$ is obtained expanding at least one leaf α of depth k as follows:*
   - *if α is open, with two nodes $\alpha \wedge \beta$ and $\alpha \wedge \neg\beta$ where β is an undecided subsentence of α.*

[1] This assumption is actually dispensable [4], but simplifies the formulation of our main definitions and results.

[2] Clearly, *Supp* is the set of leaves of F_0.

– *Otherwise, if* α *is closed, with two nodes* $\alpha \wedge \beta$ *and* $\alpha \wedge \neg\beta$*, where* $\beta \in Fm$ *is a sentence whose variables do not already occur in* $Supp \cup \{\alpha\}$*, if there are any.*

Let us notice that, when $\mathcal{F}$ is defined over a language Fm with finitely many propositional variables, the DBF may be expanded only up to a certain F_k. In what follows, given a DBF $(F_k)_{k \in \mathbb{N}}$ we will denote by $Supp_k$ the set of leaves of the forest F_k. This represent the information which is available to an agent capable of making k nested use of reasoning by cases. This information will be available to the agent for probabilistic quantification and evaluation in considering which actions to take.

3 Approximating Preferences

Our framework for preference comprises, as Savage's original one, a set of states St, a set of outcomes O, and a set of acts A. The idea is that each act $f \in A$ is a function $f\colon St \to O$.

However, we depart from Savage in various respects, in that we focus on the logical language used to represent states, rather than the more usual set-theoretic presentation.

First, we think of the set of states St as evaluations of the formulas of our logical language, of the form $v\colon Fm \to \{0, 1\}$.

Given any $f \in A$ and $S \subseteq St$ we denote by f_S the restriction of f to S. Note that a function f_S is to be interpreted as the function f when the outcomes outside S are disregarded, but it does not amount to conditioning on S, i.e. to consider the action upon the assumption that S is true, as is done e.g. in [13]. This means that, in determining, say whether f_S is preferred to g_T, both the outcomes and how likely are taken to be S, T matter.

We are now ready to reformulate some of the Savage's axioms in our setting. We focus first on those that deal with preference exclusively, without concern for their role in justifying a probabilistic representation of an agent's belief. Recall that $A \subseteq O^{St}$ and let $\succeq$ by a binary relation over A, standing for a preference between acts. Then, we require, as in Savage P1 [15]:

A1 $\succeq$ is a total *pre-order*, i.e. reflexive and transitive, over A

We then formulate a weak form of the sure-thing principle, which is closer to Savage's informal presentation [15] than to his own axiom P2.

A2 *(Sure-thing).* The following rules are satisfied:

$$\frac{f_S \succeq g_S \quad f_T \succeq g_T}{f_{S\cup T} \succeq g_{S\cup T}} \qquad \frac{f_S \succ g_S \quad f_T \succeq g_T}{f_{S\cup T} \succ g_{S\cup T}}$$

for any $S, T \subseteq Fm$ with $S \cap T = \emptyset$.

The third axiom is an adaptation of Savage's state independence P3. Before presenting it, let us recall that a set S is said to be *non-null* if there are at least two acts $f_S, f'_S \in A$ with $f_S \succ f'_S$.

A3 *(State independence).* Let $S \subseteq St$ be non-null. Then $\succeq$ satisfies the following rule:

$$\frac{f(S) = f'(St) = \{x\} \quad g(S) = g'(St) = \{y\} \quad f' \succeq g'}{f_S \succeq g_S}$$

Definition 4 (Consistent Preference Structure). *Let $A \subseteq O^{St}$ and $\succeq$ be a binary relation over A. We say that $(A, \succeq)$ is a* consistent preference structure *iff it satisfies axioms A1-A3 above.*

So far, we have only reformulated Savage's axioms, in a framework which is more congenial to our logical construction. Our key contribution is however, formalizing acts, as seen from the point of view of an agent with bounded inferential resources. Towards this purpose, we assume that the agent does not have direct access to the state space St of A, but only to some information, in a syntactic format, that she has to elaborate upon.

The actual, explicit information, provided to the agent, is here encoded by a set $Supp \subseteq Fm$. On the other hand, the information that she has to (via a reasoning effort) hypothesize about will be rendered by the set of leaves $Supp_k$ of a suitable DBF, say $\mathcal{F} = (F_k)_{k \in \mathbb{N}}$ which is built starting from $Supp$.

Now we can express what it means for an agent to access the acts via some pieces of (actual and/or hypothetical information). First, let us define

$$b_k(\varphi) = \{\alpha \in Supp_k \mid \alpha \vdash_0 \varphi\}$$

and

$$pl_k(\varphi) = \{\alpha \in Supp_k \mid \alpha \nvdash_0 \neg\varphi\}$$

in analogy with the notion of belief and plausibility function in the theory of Dempster-Shafer belief functions [16]. The set $b_k(\varphi)$ collects all the pieces of information that have been explored by the agent up to depth k, that allow her to immediately (i.e. via $\vdash_0$, without using PB) infer φ. On the other hand, $pl_k(\varphi)$ collects the pieces of information at depth k that do not immediately exclude φ.

For any $f \in A$, $f\colon St \to O$ we will denote by $f^k\colon Supp_k \to \mathcal{P}(O)$ the function associating to each piece of information $\alpha \in Supp_k$ the following subset of O:

$$f^k(\alpha) := f(\{v \in St \mid v(\varphi) = 1, \text{ for each } \varphi \text{ such that } \alpha \in pl_k(\varphi)\} \subseteq O$$

Note that a formula α is here mapped into the *set* of outcomes which are not excluded by α. This is because α, which represent a piece of information the agent *can* actually consider, need not to correspond to a state St (i.e. a logical evaluation assigning a truth value to each formula), and might not provide enough information to determine which particular outcome obtains.

Furthermore, for any $S \subseteq Supp_k$, we denote by f^k_S the restriction of f^k to S. Note that S is here taken to be a subset of *formulas* in $Supp_k$, rather than a subset of the states, i.e. of evaluations.

Definition 5 (Consistent k-Preference Structure). *Let $A \subseteq O^{St}$. We say that $(A_k, \succeq_k)$ is a* consistent k-preference structure *iff*

- A_k *contains* f^k_S *for each* $S \subseteq Supp_k$, $f \in A$
- $(A_k, \succeq_k)$ *is a consistent preference structure, i.e. it satisfies A1–A3 above.*

We are now ready to define our notion of approximating sequence.

Definition 6. *Let* $\mathcal{F} = (F_k)_{k \in \mathbb{N}}$ *be a DBF sequence, and* $A \subseteq O^{St}$. *We say that* $\mathcal{P} = (A_k, \succeq_k)_{k \in \mathbb{N}}$ *is an* approximating preference sequence *(APS, for short) iff:*

- *For each* $k \in \mathbb{N}$, $(A_k, \succeq_k)$ *is a consistent preference structure.*
- *For every* $k \in \mathbb{N}$, *and every* $\varphi, \psi \in Fm$, $f, g \in Supp$, *we have that* $f^k \succeq_k g^k$ *implies* $f^{k'} \succeq_{k'} g^{k'}$ *for every* $k' \geq k$.

The second condition says that, as k increases, the agent can refine, but cannot revise previously determined preferences. Let us test now our notion of APS with the well-known examples of Ellsberg and Allais. To ease notation, in the following we will often slightly abuse the notation, writing directly $f \succeq_k g$ instead of $f^k \succeq_k g^k$.

Example 2 (Ellsberg). Suppose that an agent is presented an urn filled with balls, and is provided the information that 2/3 of the balls are either yellow or blue ($Y \vee B$), and the remaining 1/3 are red (R). A ball will be extracted from the urn and an agent is confronted with a choice between acts f, g, h, j. The following table summarizes the setup in the standard Savage framework, where states are represented in the columns, the available acts in the rows, and the cells contain the monetary outcome, say in euros.

Table 2. Ellsberg's one urn scenario.

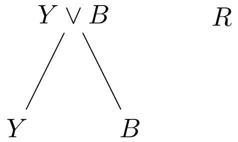
	R	Y	B
f	100	0	0
g	0	100	0
h	100	0	100
j	0	100	100

Ellsberg [10] points out that the strict preferences $f \succ g$ and $j \succ h$ are plausible: agents will typically prefer, ceteris paribus, a bet whose states they can quantify probabilistically (R and $Y \vee B$ for the acts f and j) over one where this is not the case (Y and B for the acts g and j). In other words, they will display a form of *ambiguity aversion* [12].

On the other hand, these preferences are in violation of Savage STP. Indeed, if we ignore what happens in case a blue ball (B) is picked (i.e. we ignore the third column in Table 2), and we assume that the preference for a payoff of 100 euros is independent of the state in which it occurs, the agent should be indifferent between acts f and h, and g and j. Furthermore, both, f and g, and h and j give the same payoff for B, i.e. 0 and 100, respectively. According to the STP then, a preference for f over h dictates a preference for g over j, in contrast to Ellsberg's preferences.

Let us now formalize the example in our setting. We take a finite language over the variables $\{Y, B, R\}$ which stand for the event that a yellow, blue, red ball is picked, respectively. We denote by γ the sentence expressing the fact that Y, B, R are mutually exclusive and jointly exhaustive. We build a DBF and an APS as follows. We let $Supp = \{(Y \vee B) \wedge \gamma, R \wedge \gamma\}$, since those are the formulas upon which the agent is provided probabilistic information, and $A = \{f, g, h, j\}$. It is easy to show that for any such formula $\alpha \in Supp$ we have $pl_0(\alpha) = \{\alpha\}$. The acts f, g, h, j are again defined as in Table 2. Assume that $f \succ_0 g$ and $j \succ_0 h$. We may consider a decomposition of such preferences only via the formulas in *Supp*. We have (omitting the formula γ, for simplicity): $g_{Y \vee B} \succ_0 f_{Y \vee B}$, $f_R \succ_0 g_R$, $j_{Y \vee B} \succ_0 h_{Y \vee B}$, and $h_R \succ_0 j_R$. These preferences, together with $f \succ_0 g$ and $j \succ_0 h$, do not contradict axiom A2, i.e. our reformulated version of the Sure-thing principle. Note that, since $Y \vee R$ and B are not formulas of $Supp_k$, the functions say $f_{Y \vee R}, h_{Y \vee R}, g_{Y \vee R}, j_{Y \vee R}$ and f_B, h_B, g_B, j_B are not defined.

Now, let us consider the expansion of *Supp* to a 1-depth forest F_1, and the corresponding 1-depth preference structure over $Supp_1$. Notice that the node $R \wedge \gamma$ in *Supp* is already closed, and thus need not be expanded. We expand instead the open node $(Y \vee B) \wedge \gamma$ as follows (we omit γ for simplicity):



Consider the preference structure $(A_1, \succeq_1)$. With a little abuse of notation, since $((Y \vee B) \wedge \gamma \wedge Y) \vdash_0 Y$, $((Y \vee B) \wedge \gamma \wedge \neg Y) \vdash_0 B$ and $R \wedge \gamma \vdash_0 R$, we just write the formula on the right Y, B, R instead of the corresponding formula on the left, which belongs to $Supp_1$.

Note that, at depth 1, the preferences $f \succ_1 g$ and $j \succ_1 h$ are not allowed by Definition 5. By state independence, we have indeed that $f_{\{Y\}} \approx_1 h_{\{Y\}}$, $f_{\{R\}} \approx_1 h_{\{R\}}$ and $g_{\{Y\}} \approx_1 j_{\{Y\}}$, $g\{R\} \approx_1 j_{\{R\}}$. On the other hand, we have $f_{\{B\}} \approx_1 g_{\{B\}}$, and $h_{\{B\}} \approx_1 j_{\{B\}}$, while $j_{\{B\}} \succ f_{\{B\}}$.

Now, let us further assume that $f_{\{Y\} \cup \{R\}} \succ_1 g_{\{Y\} \cup \{R\}}$. By the previous equivalences, we may use A2 to get $h_{\{Y\} \cup \{R\}} \succ j_{\{Y\} \cup \{R\}}$. By the latter, since we also have $h_{\{B\}} \approx_1 j_{\{B\}}$ we may use A2 to obtain $h \succ_1 j$, which is contrary to the initial assumption $j \succ_1 h$.

Let us now assume $g_{\{Y\} \cup \{R\}} \succeq_1 f_{\{Y\} \cup \{R\}}$. Since $f_B \approx_1 g_B$, by state independence, we obtain by A2, $g = g_{\{Y\} \cup \{R\} \cup \{B\}} \succeq_1 f_{\{Y\} \cup \{R\} \cup \{B\}} = f$, again contradicting the initial assumption that $f \succ_1 g$. In both cases we derived a contradiction with one of the assumptions $f \succ_1 g$ and $j \succ_1 h$.

Example 3 (Allais). Assume you have an urn containing balls numbered from 1 to 100, and a ball will be extracted from the urn. You are offered a choice between the following acts, which are represented in the following table.

Table 3. Allais.

	1	2–10	11–100
f	100	100	100
g	0	500	100
f'	100	100	0
g'	0	500	0

Allais deems the strict preferences $f \succ g$ and $g' \succ f'$ plausible, although they conflict with the sure-thing principle. Indeed, the pairs of acts f and g, and f' and g' have the same outcome, in case balls 11-100 are extracted, namely 100 for the first pair, and 0, for the second. By the sure-thing principle, since the acts f and f', and g and g' have the same outcomes for each extracted ball, f can be preferred to g, if and only f' is preferred to g'.

We formalize this scenario in our setting, building a DBF and an APS. It suffices to consider a finite language over three variables, namely $\{p_1, p_{2-10}, p_{11-100}\}$, standing for the numbers on the extracted ball. We let $Supp = \{\gamma\}$ where γ encodes the fact that $p_1, p_{2-10}, p_{11-100}$ are mutually exclusive and jointly exhaustive. We further let $A = \{f, g, f', g'\}$, where the acts are defined as in Table 3. At depth 0, we may only compare $f_\gamma, g_\gamma, g'_\gamma, f'_\gamma$, since $Supp = \{c\}$. Hence, we may have $f \succeq_0 g$ and $g' \succeq_0 f'$, since no application of A2 can be performed. At depth 1, we replace $Supp$ with $Supp_1 = \{\gamma \wedge \neg p_{11-100}, \gamma \wedge p_{11-100}\}$. We omit γ in the following for simplicity. We have $f_{\neg p_{11-100}} \approx_1 f'_{\neg p_{11-100}}$, $g_{\{\neg p_{11-100}\}} \approx_1 g'_{\{\neg p_{11-100}\}}$, and on the other hand $f_{p_{11-100}} \approx_1 g_{p_{11-100}}$ and $f'_{p_{11-100}} \approx_1 g'_{p_{11-100}}$. By A2 we immediately get that $f_{Supp_1} \succeq_1 g_{Supp_1}$ iff $f'_{Supp_1} \succeq_1 g'_{Supp_1}$, contrary to the Allais' preferences.

Finally, we address Example 1 in our formal setting.

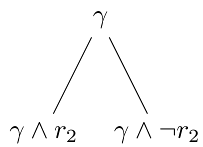
Example 1 (continued). We denote:[3] by p_{in} the assertion "$x_i = n''$; by q_{ij}, the assertion "$x_i = x_j + 1''$ and finally by r_i the assertion "the ith extracted ball is red". The initial information provided to the agent is $Supp = \{\gamma\}$, where by γ we denote the formula $r_1 \wedge \neg r_3 \wedge q_{12} \wedge q_{23} \wedge \bigvee_{k=1}^{100} p_{1k}$. The formula δ in Example 1 is encoded instead as:

$$\bigvee_{\substack{i,j \in \{1,2,3\} \\ i \neq j}} r_i \wedge \neg r_j \wedge q_{ij}.$$

We take $A = \{h, h'\}$, where h, h' are defined as in Example 1, with $h(\gamma) = \{100\}$ and $h(\neg\gamma) = \{0\}$, and $h'(\delta) = \{110\}$, $h'(\neg\delta) = \{0\}$. Now, in A_0 we may compare h^0 and h'^0, which both have $Supp = \{\gamma\}$ as their domain. We have then $h^0(\gamma) = \{100\}$ and $h'^0(\gamma) = h'(\{\delta, \neg\delta\}) = \{110, 0\}$ since $\gamma \not\vdash_0 \delta$. Hence we may still allow

[3] We use a propositional language, to fit the simple general framework put forward in this work, although we might have a more compact representation in a first-order language.

$h \succ_0 h'$. On the other hand, if we consider the 1-depth forest (actually, tree) expanding $Supp = \{\gamma\}$ as follows:



we now have that both $h'^1(\gamma \wedge r_2) = \{110\}$ and $h'^1(\gamma \wedge \neg r_2) = \{110\}$, since $\gamma \wedge r_2 \vdash_0 \delta$ and $\gamma \wedge \neg r_2 \vdash_0 \delta$. Hence h'^1 is constantly equal to 110. On the other hand h^1 is still constantly equal to 100, and assuming that 110 is preferred to 100, we may only have $h' \succeq_1 h$, by state independence.

4 Qualitative Probability and Representation

So far, we have build up the general framework and illustrated how it takes into account various alleged counterexamples, and criticisms of Savage's approach. In particular, our setting shows that a form of idealization is at play in Savage's setting, in essentially disregarding the cost of reasoning by case.

This does not preclude to obtain as a limit, idealized case, Savage's elegant mathematical result, in our framework. Let us recall that one of the main advantages of Savage's framework is its representation theorem for expected utility, which is obtained on the basis of his axioms on preferences among acts. While we are still not able to recover the full representation of expected utility in the limit, in our setting, we will focus here on an important intermediate step towards this result, which has an independent foundational interest.

Let us recall that, on the way to his representation theorem, Savage first manages to obtain a measure of probability, only on the basis of preferences among acts. This is done in two steps: first he derives, from the preference of an agent, an ordering reflecting how likely the agent finds the events of interest, i.e. a *qualitative probability*. Subsequently, he extracts from this relational structure a unique numerical probability representing it.

Let us now recall the notion of qualitative probability over arbitrary boolean algebras, and that of representability, and adapt them to our setting.

Definition 7 (Qualitative probability). *Let* $\mathcal{B} = (B, \sqsubseteq, \wedge, \vee, \neg, \bot, \top)$ *be a boolean algebra.* $(\mathcal{B}, \trianglerighteq)$ *is a* qualitative probability *if*

1. $\trianglerighteq$ *is a total preorder over* $\mathcal{B}$*;*
2. $\top \triangleright \bot$*;*
3. *if* $\alpha \sqsupseteq \beta$ *then* $\alpha \trianglerighteq \beta$ *and*
4. *if* $\alpha \wedge \gamma = \bot$ *and* $\beta \wedge \gamma = \bot$*, then* $\alpha \trianglerighteq \beta$ *if and only if* $\alpha \vee \gamma \trianglerighteq \beta \vee \gamma$*.*

Since our sequences are built syntactically, we will use here a different, syntactic definition of qualitative probability.

Definition 8 (synctactic qualitative probability). *Let Fm be the set of formulas over the language $\mathcal{L}$. $(Fm, \trianglerighteq)$ is a* (syntactic) qualitative probability *if*

1. $\trianglerighteq$ *is a total preorder over Fm;*
2. $\top \triangleright \bot$;
3. *if* $\beta \vdash \alpha$ *then* $\alpha \trianglerighteq \beta$ *and*
4. *if* $\alpha \wedge \gamma \vdash \bot$ *and* $\beta \wedge \gamma \vdash \bot$ *then*

$$\alpha \trianglerighteq \beta \text{ if and only if } \alpha \vee \gamma \trianglerighteq \beta \vee \gamma.$$

The two notions are essentially equivalent. Indeed, if we are given a (syntactic) qualitative probability $(Fm, \trianglerighteq)$, we may just define a qualitative probability by quotienting over the logically equivalent formulas, i.e. building the Lindenbaum-Tarski algebra and suitably adapting the $\trianglerighteq$ relation to the equivalence classes. Let us now recall the following, see e.g. [15].

Definition 9 ((Almost) Representability). *A qualitative probability $(\mathcal{B}, \trianglerighteq)$ is said to be*

- representable *if there exists a unique*[4] *finitely additive probability P such that* $\alpha \trianglerighteq \beta$ *iff* $P(\alpha) \geq P(\beta)$
- almost representable*, if there exists a unique finitely additive probability P such that* $\alpha \trianglerighteq \beta$ *implies* $P(\alpha) \geq P(\beta)$.

Savage considers in his system a specific axiom P4 for the purpose of extracting a qualitative probability from preference, and a further axiom P6 for the purpose of representability. In our framework, we obtain qualitative probabilities and representability via a slightly different route, inspired by the reformulation of P4 in [6].

First, we will define a sequence of *comparative beliefs*, determined by an APS.

Definition 10. *Let $\mathcal{F} = (F_k)_{k \in \mathbb{N}}$ be a DBF and $(A_k, \succeq_k)_{k \in \mathbb{N}}$ be an APS. We call* comparative plausibility $\trianglerighteq_k$ determined by $\succeq_k$*, the relation $\trianglerighteq_k$ defined, for any $\varphi, \psi \in Fm$ by:*

- $\varphi \trianglerighteq_k \psi$ *if* $f^k_\varphi \succeq_k g^k_\psi$*, for each* $f^k, g^k \in Supp_k$ *such that* $f^k(\varphi) = g^k(\psi) = \{x\}$ *for some* $x \in O$.
- $\varphi \trianglerighteq_k \psi$ *if* $pl_k(\varphi) \supseteq pl_k(\psi)$.
- $\top \triangleright_k \bot$

The idea is that, when we consider acts that have the same outcome, over different pieces of information, the preferences of an agent for one act over the other, only reflects how likely she finds the piece of information to occur. More concretely, if an agent prefers a bet giving her 5 euros if tomorrow it rains, to a bet giving her 5 euros if tomorrow it will be sunny, this can only mean (if she is rational) that she finds rainy weather more likely than sunny weather.

[4] Uniqueness is typically nor requested in the definition of representability and almost representability in the literature.

Note that the definition ensures that $\trianglerighteq_k$ is not empty, hence in particular it encodes Savage's axiom ($P5$).

We now give conditions on an APS, to obtain from the sequences of $\trianglerighteq_k$, a qualitative probability *in the limit*. Before that, we recall, adapting from [4] what we mean by limit.

Definition 11 (Limit structures). *Take a DBF and let $\mathcal{F} = (Supp_k, \trianglerighteq_k)_{k\in\mathbb{N}}$ be a sequence of relational structures, where each $\trianglerighteq_k$ is a binary relation over Fm. We say that the structure $(Fm, \trianglerighteq)$ is* the limit of $\mathcal{F}$*, where*

$$\varphi \trianglerighteq \psi \text{ iff there is a } k \text{ such that } \varphi \trianglerighteq_n \psi, \text{ for every } n \geq k.$$

Definition 12. *We say that an APS $\mathcal{P} = (A_k, \succeq_k)_{k\in\mathbb{N}}$ over a DBF $\mathcal{F} = \{F_k\}_{k\in\mathbb{N}}$ is:*

- Belief-determining *iff:*
  - *For any $\varphi, \psi \in Fm$ there exists a $k \in \mathbb{N}$ such that either $\varphi \trianglerighteq_k \psi$ or $\psi \trianglerighteq_k \varphi$.*
- Refinable *if whenever $\alpha \trianglerighteq_k \beta$ for some $\alpha, \beta \in Supp_k$ and $k \in \mathbb{N}$, there is a $k' \geq k$ such that*

$$\beta \triangleright_{k'} \gamma \text{ for every } \gamma \in Supp_{k'} \text{ that is a descendent of } \alpha.$$

- Coverable *if whenever $\alpha \triangleright_k \beta$ for some $\alpha, \beta \in Supp_k$ and $k \in \mathbb{N}$, there is a $k' \geq k$ and $\gamma \in Supp_{k'}$ such that $\gamma \wedge \alpha \vdash \bot$ and*

$$\alpha \vee \gamma \bowtie_{k'} \beta$$

The condition of being belief-determining is our reformulation of axiom P4 in Savage, which is here considered as an axiom of a whole APS, rather than of each Consistent k-Preference Structure, as we did instead for A1–A3. By this condition, indeed, $\trianglerighteq_k$ determines a total order in the limit.

We are now ready to provide our main result.

Theorem 1. *Let $\mathcal{P}$ be an APS over a DBF $\mathcal{F}$ with $Supp = \{*\}$. If $\mathcal{P}$ is belief-determining, then the limit $(Fm, \trianglerighteq)$ of $(F_k, \trianglerighteq_k)_{k\in\mathbb{N}}$ is a qualitative probability.*

Proof. Let us start by showing that, if $\psi \vdash \varphi$, then $\varphi \trianglerighteq \psi$. From $\psi \vdash \varphi$, we get $\neg\varphi \vdash \neg\psi$. We thus have a derivation of $\neg\psi$ from $\neg\varphi$, by using the rules of $\vdash_0$ and applications of PB. Let $k \in \mathbb{N}$ be such that for any $n \geq k$, the set $Supp_n$ collects all the premises of the applications of PB in the proof of $\neg\psi$ from $\neg\varphi$. Hence, for each $\alpha \in Supp_n$, if $\alpha \vdash_0 \neg\varphi$, then $\alpha \vdash_0 \neg\psi$, that is, if $\alpha \nvdash_0 \neg\psi$, then $\alpha \nvdash_0 \neg\varphi$. Hence $pl_n(\varphi) \supseteq_n pl_n(\psi)$, for $n \geq k$. This entails, by Definition 10, $\varphi \trianglerighteq_n \psi$, for each $n \geq k$, hence $\varphi \trianglerighteq \psi$.

We now show that the relation is total. Take $\varphi, \psi \in Fm$. Now, since $\mathcal{P}$ is belief determining, there is a k such that $\varphi \trianglerighteq_k \psi$ or $\psi \trianglerighteq_k \varphi$. Assume the first is the case. Since $\mathcal{P}$ is an APS, we will also have that, for any $n \geq k$, $\varphi \trianglerighteq_n \psi$, hence in particular $\varphi \trianglerighteq \psi$.

Transitivity and reflexivity are immediate, since they follow by A1 for $\succeq_k$, and the fact that $\mathcal{P}$ is an APS.

As for additivity, suppose that $\varphi \wedge \chi \vdash \bot$ and $\psi \wedge \chi \vdash \bot$. We will show that $\varphi \trianglerighteq \psi$ iff $\varphi \vee \gamma \trianglerighteq \psi \vee \gamma$. Let k be such that each $\alpha \in Supp_k$ is closed. We have that $\varphi \vee \chi \trianglerighteq_k \psi \vee \chi$ iff $\varphi \trianglerighteq_k \psi$ (adapting the proof of Lemma 11(5) in [4]). By the definition of $\trianglerighteq_k$, this means that for each f, g such that $f^k(\varphi \vee \xi) = g^k(\varphi \vee \psi) = \{x\}$ we have $f_{\varphi \vee \xi} \succeq_k g_{\psi \vee \xi}$. On the other hand, by the reflexivity of $\succeq_k$ (A1), we have $f_\xi \succeq_k f_\xi$ and $g_\xi \succeq_k g_\xi$. Hence, by A2 $f_{\{\varphi\} \cup \{\xi\}} \succeq_k f_{\{\psi\} \cup \{\xi\}}$ iff $f_\varphi \succeq_k g_\psi$. But the latter amounts at saying that $\varphi \trianglerighteq_k \psi$, and the same will hold for any $n \geq k$. Hence we have finally obtained $\varphi \trianglerighteq \psi$ iff $\varphi \vee \xi \trianglerighteq \psi \vee \xi$.

Finally, adapting from [4], we have that, under the refinability and coverability conditions described above, an APS determines a (almost) representable qualitative probability.

Corollary 1. *Let $\mathcal{P}$ be a belief-determining APS.*

- *If $\mathcal{P}$ is refinable, then its limit is almost representable, in the case $\mathcal{A}_\mathcal{L}$ is infinite.*
- *If $\mathcal{P}$ is coverable then its limit is representable, in the case $\mathcal{A}_\mathcal{L}$ is finite.*

Proof. Follows from Theorem 1, and Theorem 20 and 22 in [4].

5 Conclusion

We have introduced a logic-based framework for preference, which approximates Savage's framework, on the basis of the bounded use of hypothetical information. Our approach accommodates in a unified way various traditional challenges to Savage, in particular concerning the Sure-thing principle. Despite their differences, in all the examples considered, we have found indeed a similar pattern: some preferences may be accepted at the bottom level of our sequence, i.e. $\succeq_0$, but they turn out to be inconsistent with Savage-style axioms, when considering $\succeq_k$ for $k > 0$, i.e. when suitable hypothetical information is taken into account. Since DBBLs are computationally tractable, a further natural direction of research for our work is in the computational complexity issues related with the reasoning with the resulting measures of comparative probability. In particular, we aim to compare our setting with other approaches to decision theory, which are logically (in particular, syntactically) and computationally inspired, such as that pursued in [5].

Future work will provide suitable representation theorems for preferences in our framework, in terms of generalized expected utility, both at each level of the approximating sequence, and in the limit. This will be compared with the literature on decision-making under uncertainty, based on weakenings of axioms in the Anscombe-Aumann framework [11]. We further plan to consider logical systems where the preference relation $\succeq_k$ is taken to be part of the language, and investigate their properties, with the aim of obtaining tractable logics of preference.

References

1. Allais, M.: Le comportemement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine. Econometrica **21**(4), 503–546 (1953)
2. Aumann, R.J., Hart, S., Perry, M.: Conditioning and the sure-thing principle, pp. 1–10 (2005)
3. Baldi, P., Hosni, H.: Tractable approximations of probability. J. Logic Comput. **33**, 599–622 (2022)
4. Baldi, P., Hosni, H.: Logical approximations of qualitative probability. In: ISIPTA. Proceedings of Machine Learning Research, vol. 147, pp. 12–21. PMLR (2021)
5. Bjorndahl, A., Halpern, J.Y.: Language-based decisions. Electron. Proc. Theor. Comput. Sci. **335**, 55–67 (2021)
6. Bradley, R.: Decision Theory with a Human Face. Cambridge University Press, Cambridge (2017)
7. D'Agostino, M.: An informational view of classical logic. Theor. Comput. Sci. **606**, 79–97 (2015)
8. D'Agostino, M., Finger, M., Gabbay, D.: Semantics and proof-theory of depth bounded Boolean logics. Theor. Comput. Sci. **480**, 43–68 (2013)
9. D'Agostino, M., Floridi, L.: The enduring scandal of deduction: is propositional logic really uninformative? Synthese **167**, 271–315 (2009)
10. Ellsberg, D.: Risk, ambiguity, and the savage axioms. Quart. J. Econ. **75**(4), 643–669 (1961)
11. Gilboa, I.: Theory of Decision Under Uncertainty. Cambridge University Press, Cambridge (2009)
12. Gilboa, I., Marinacci, M.: Ambiguity and the Bayesian paradigm. In: Arló-Costa, H., Hendricks, V.F., van Benthem, J. (eds.) Readings in Formal Epistemology. SGTP, vol. 1, pp. 385–439. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-20451-2_21
13. Kranz, D., Luce, R.D., Suppes, P., Tversky, A.: Foundations of Measurement, vol. 1. Academic Press, New York (1971)
14. Pearl, J.: The sure-thing principle. J. Causal Inference **4**(1), 81–86 (2016)
15. Savage, L.J.: The Foundations of Statistics, 2nd edn. Dover (1972)
16. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
17. Sipser, M.: Introduction to the Theory of Computation. PWS Publishing Company (1997)
18. Stanovich, K.: What Intelligence Tests Miss: The Psychology of Rational Thought. Yale University Press (2009)