THE MODAL LOGIC OF PROBABILITY
Extended Abstract

Aviad Heifetz*

Philippe Mongin†‡

1 Introduction

The present paper offers a novel axiomatization of the probability concept in terms of modal logic. The structures we axiomatize consist of a measurable space of possible worlds, and for each possible world, a probability measure on the space and a valuation function. Roughly speaking, these structures can be seen as probabilistic refinements of the familiar Kripke structures of modal logic. Leaving aside measurability restrictions, the difference between the two kinds of structures is simply that in a *probability structure* each world is mapped to a probability measure instead of a set. Conversely, a Kripke structure can be seen as an impoverishment of a probability structure, in which only supports (i.e., minimal closed sets of probability one) are considered. Similarly to Kripke structures, probability structures admit of various conceptual interpretations, but this paper was motivated by earlier work in epistemic logic and the foundations of decision theory and game theory, so we will be exclusively concerned here with the interpretation of probability as a measure of subjective belief.

For simplicity, our formalism and main results are stated for a single individual, but they can be trivially extended to the multi-agent case. When each agent has a mapping from possible worlds to probabilities, the structures become what game theorists call a *type space*. The latter concept was introduced by Harsanyi (1967-68) in his pioneering attempt to build a general set-up for games with incomplete information. Type spaces have the property of summarizing at once the players' uncertainties about the game

*The School of Economics, Tel Aviv University, aviad@econ.tau.ac.il

†THEMA, CNRS, Université de Cergy-Pontoise, mongin@u-cergy.fr

‡This work is part of the research in the TMR network "Cooperation and Information", ERB FMRX CT 0055. The authors are grateful for financial support from the network.

and the other players' uncertainties. Formally, a type space involves a measurable state space, and for each state and each agent, a probability measure on the space. It also includes a specification of the objective parameters of the game in each state, which is the analogue of a valuation function.

The generality and wide applicability of probability structures make them a natural candidate for logical investigations. Fagin and Halpern (1994) and Fagin, Halpern and Megiddo (1990) have axiomatized probability structures with a very rich syntax, which can express not only probabilities of formulas, but also probabilities of *linear combinations of formulas.*

Here we follow a completely different route. Our more restricted language does not make it possible to express directly addition or scalar multiplication of formulas. Rather, it is a restricted extension of the propositional modal syntax. All epistemic features are captured by belief operators L_α for rational $\alpha \in [0,1]$, to be interpreted as "the probability is at least α". Had we restricted α to be only 0 or 1, the system presented here would collapse into the familiar modal system KD (see, e.g., Chellas 1980).

This syntax with indexed operators was suggested by Aumann (1995, section 11). The set of axioms he states there is sound with respect to probability structures. We show, however, that it is not complete. The way we propose to complete Aumann's system might seem roundabout at first sight, and indeed it remains an open problem whether there is a simpler complete system. Nevertheless, the "complicated" extra axiom that we need is closely related to a technical condition introduced to resolve two fundamental issues in probability theory.

The first is the existence of a probability compatible with a qualitative ordering of events . De Finetti stated a set of simple necessary conditions on the ordering to be compatible with a probability, and asked whether they are sufficient. Kraft, Pratt and Seidenberg (1959) constructed a counterexample, and introduced the missing condition for sufficiency – a condition to which our extra axiom is very close in spirit.

The second issue is how to characterize the pairs of a super-additive lower probability and a sub-additive upper probability which can be separated by an (additive) probability. (For instance, Walley (1981) and Papamarcou and Fine (1986) showed this cannot always be done.) Suppes and Zanotti (1989, theorem 1) introduced a necessary and sufficient condition, which is again similar to the one we employ. In all of these cases, the condition is needed to use some version of the separation theorem or the theorem of the alternative in convex analysis.

Like in ordinary modal logic, our language allows only for *finite* conjunctions and disjunctions. This approach leads to a well-recognized difficulty in investigating the probability calculus (on the connection with infinitary first-order logic see Gaifmann (1964) and Scott and Krauss (1966)). This is the problem of "non-Archimedianity". For example, the set of formulas Σ which says that the probability of φ is at least $\frac{1}{2} - r$ for every rational r, is consistent with the formula ψ which says that the probability of φ is strictly smaller than $\frac{1}{2}$, because every finite subset of $\Sigma \cup \{\psi\}$ is consistent. However, there is no real number for the probability of φ which would be compatible with the whole set $\Sigma \cup \{\psi\}$. When infinite conjunctions and disjunctions are permitted, Σ could be made to imply the negation of ψ, and thus avoid the problem. But in a finitary logic this cannot be done, so there is no hope to have *strong* completeness of the system[1]. Put differently, the canonical space of maximally consistent sets of formulas cannot be endowed with a probability structure compatible with the formulas that build the states, because in those states that contain the formulas of $\Sigma \cup \{\psi\}$, there is no suitable probability for the set of states containing φ. We can still strive, however, to find a complete system, in which every tautology – i.e., semantic truth in the family of probability structures, spelled out in the finitary language – is a theorem of the system, i.e. is provable from its axioms.

We circumvent the difficulty of non-Archimedianity by employing a device which has led to successful completeness proofs of finitary axiomatizations of *common belief.* The idea is to choose suitable *filtrations* of the full language to sub-languages with finitely many formulas.

2 The formal system and the determination theorem

The formal language $\mathcal{L}$ in this paper is built in the familiar way from the following components: A set $\mathcal{P}$ of propositional variables, the connectives $\neg$ and $\wedge$, from which the other connectives $\vee$, $\rightarrow$ and $\leftrightarrow$ are defined as usual, and the modal operators L_α for any rational α in $[0, 1]$, with the intended meaning "the probability is at least α". The operator M_α – "the probability

[1] That is, we cannot hope to have a finitary system for probability in which if ψ holds whenever a (possibly infinite) set of formulas Σ holds in a probability structure, then Σ proves ψ.

is at most α" – is an abbreviation defined by

$$M_\alpha\varphi \leftrightarrow L_{1-\alpha}\neg\varphi,$$

and the operator E_α – "the probability is equal to α" is defined by

$$E_\alpha\varphi \leftrightarrow M_\alpha\varphi \wedge L_\alpha\varphi.$$

The operator S_α – "the probability is strictly smaller than α" is defined by $S_a\varphi \leftrightarrow \neg L_\alpha\varphi$, and similarly the operator G_α – "the probability is strictly greater than α" is defined by $G_a\varphi \leftrightarrow \neg M_\alpha\varphi$.

The space $\mathcal{M}$ of *probability structures* that we aim to axiomatize has a typical element

$$m = \langle \Omega, \mathcal{A}, P, v\rangle$$

where Ω is a non-empty set; $\mathcal{A}$ is a σ-field of events – subsets of Ω; P is a measurable mapping from Ω to the space $\Delta(\Omega, \mathcal{A})$ of probability measures, which is endowed with the σ-field generated by the sets

$$\{\mu \in \Delta(\Omega, \mathcal{A}) : \ \mu(E) \geq \alpha\} \qquad \text{for all } E \in \mathcal{A} \text{ and rational } \alpha \in [0,1]$$

and v is a mapping from $\Omega \times \mathcal{P}$ to $\{0,1\}$, such that $v(\cdot, p)$ is measurable for every $p \in \mathcal{P}$.

The validation clauses of our logic are stated inductively in the usual way for the propositional connectives, and as follows for the modal operators L_α (resp. M_α, E_α, S_α, G_α):

$$m, \omega \models L_\alpha\varphi \quad \text{iff} \quad P(\omega)([\varphi]) \geq \alpha \quad (\text{resp. } \leq \alpha, = \alpha, < \alpha, > \alpha)$$

where

$$[\varphi] = \{\omega \in \Omega : m, \omega \models \varphi\}.$$

We use the familiar abbreviations, $m \models \varphi$ for $[\forall\omega \in \Omega,\ m, \omega \models \varphi]$, and $\mathcal{M} \models \varphi$ for $[\forall m \in \mathcal{M},\ m \models \varphi]$.

A starting point for axiomatization is the following system $\mathcal{A}$, which we have adapted from Aumann (1995, section 11). (The symbol $\vdash$ denotes the inference relation of the system, and $\top$ and $\bot$ abbreviate $\varphi \vee \neg\varphi$, $\varphi \wedge \neg\varphi$, respectively.)

Any axiomatization of the propositional calculus (A0)

$$L_0\varphi \tag{A1}$$

$$L_\alpha \top \tag{A2}$$

$$L_\alpha(\varphi \wedge \psi) \wedge L_\beta(\varphi \wedge \neg\psi) \rightarrow L_{\alpha+\beta}\varphi, \qquad \alpha + \beta \leq 1 \tag{A3}$$

$$S_\alpha(\varphi \wedge \psi) \wedge S_\beta(\varphi \wedge \neg\psi) \rightarrow S_{\alpha+\beta}\varphi, \qquad \alpha + \beta \leq 1 \tag{A4}$$

$$L_\alpha\varphi \rightarrow S_\beta\neg\varphi \qquad \alpha + \beta > 1 \tag{A5}$$

$$\text{If } \ \vdash \varphi \leftrightarrow \psi \ \text{ then } \ \vdash L_\alpha\varphi \leftrightarrow L_\alpha\psi \tag{A6}$$

Proposition 1: *From $\mathcal{A}$ the following axiom and inference rule schemata can be derived:*

$$\text{If } \ \vdash \varphi \rightarrow \psi \ \text{ then } \ \vdash L_\alpha\varphi \rightarrow L_\alpha\psi \tag{A6+}$$

$$L_\alpha\varphi \rightarrow L_\beta\varphi \qquad \beta < \alpha \tag{A7}$$

$$L_\alpha\varphi \rightarrow G_\beta\varphi \qquad \beta < \alpha \tag{A7+}$$

$$S_\alpha\varphi \rightarrow M_\alpha\varphi \tag{A8}$$

$$L_\alpha(\varphi \wedge \psi) \wedge G_\beta(\varphi \wedge \neg\psi) \rightarrow G_{\alpha+\beta}\varphi, \qquad \alpha + \beta \leq 1 \tag{A9}$$

Aumann's (1995) axiom system actually consists of (A0)-(A5), (A6+) and (A7). The schemata (A3), (A4) and (A9) translate into the formal language inequalities of the probability calculus for disjoint events. For disjoint A and B these are:

(i) $\mu(A) \geq \alpha,\ \mu(B) \geq \beta \ \Rightarrow \ \mu(A \cup B) \geq \alpha + \beta$

(ii) $\mu(A) \geq \alpha,\ \mu(B) > \beta \ \Rightarrow \ \mu(A \cup B) > \alpha + \beta$

(iii) $\mu(A) \leq \alpha,\ \mu(B) \leq \beta \ \Rightarrow \ \mu(A \cup B) \leq \alpha + \beta$

(iv) $\mu(A) \leq \alpha,\ \mu(B) < \beta \ \Rightarrow \ \mu(A \cup B) < \alpha + \beta$

In the system $\mathcal{A}$, the schemata (A3) and (A9) express (i) and (ii). The schema (A4) expresses only the following form, which is *weaker* than (iv):

(v) $\mu(A) < \alpha,\, \mu(B) < \beta \;\Rightarrow\; \mu(A \cup B) < \alpha + \beta$

Proposition 2: *The system $\mathcal{A}$ does not imply the following schemata, (which express (iii) and (iv)):*

$$M_\alpha(\varphi \wedge \psi) \wedge M_\beta(\varphi \wedge \neg\psi) \rightarrow M_{\alpha+\beta}\varphi, \qquad \alpha + \beta \leq 1 \tag{A10}$$

$$M_\alpha(\varphi \wedge \psi) \wedge S_\beta(\varphi \wedge \neg\psi) \rightarrow S_{\alpha+\beta}\varphi, \qquad \alpha + \beta \leq 1 \tag{A11}$$

Since $\mathcal{M} \models$ (A10), (A11), the system $\mathcal{A}$ is not complete with respect to $\mathcal{M}$.

It is not hard to show that $\mathcal{A}$+(A10) implies (A11). Therefore, it would be enlightening to know whether the system $\mathcal{A}$+(A10) is complete with respect to the family of probability structures. This question remains an open problem for the time being. We do suggest, however, to substitute (A3) with a unique, more elaborate schema, which will imply at once (A3), (A9),(A10) and (A11), and which will complete the system. Before we introduce the axiom formally, we explain its intuitive meaning.

Recall that a probability measure μ on a space Ω defines the integral functional on the collection of characteristic functions (of measurable sets), and hence also on the semi-group of finite sums of such characteristic functions. In particular, if a function f in this semi-group can be written as a sum of characteristic functions in two different ways, then the two ways of calculating the integral give the same result. Explicitly, f is the sum of the characteristic functions of $E_1, \ldots, E_m$, and also of $F_1, \ldots, F_n$, if and only if the points that belong to at least one of $E_1, \ldots, E_m$ belong to at least one $F_1, \ldots, F_n$ and vice versa, and similarly for the points that belong to at least two sets, three sets, etc. Let us denote by $E^{(k)}$ the points that appear in at least k of the sets $E_1, \ldots, E_m$, and by $F^{(k)}$ the points that appear in at least k of the sets $F_1, \ldots, F_n$, i.e.

$$E^{(k)} = \bigcup_{1 \leq \ell_1 < \ldots < \ell_k \leq m} (E_{\ell_1} \cap \ldots \cap E_{\ell_k})$$

$$F^{(k)} = \bigcup_{1 \leq \ell_1 < \ldots < \ell_k \leq n} (F_{\ell_1} \cap \ldots \cap F_{\ell_k})$$

We clearly have the following property, that we denote by $(C1)$: If[2]

$$E^{(k)} = F^{(k)} \text{ for } 1 \leq k \leq \max(m,n)$$

then

$$\mu(E_i) \geq \alpha_i \text{ for } i = 1, \dots m$$

and

$$\mu(F_j) \leq \beta_j \text{ for } j = 1, \dots n-1,$$

entail that

$$\mu(F_n) \geq (\alpha_1 + \dots + \alpha_{m)} - (\beta_1 + \dots \beta_{n-1}),$$

– otherwise the integral of f would not be well defined.

The axiom that we propose is a syntactical rendering of property (C1). If $(\varphi_{1,} \dots, \varphi_m)$ is a collection of formulas, we abbreviate by $\varphi^{(k)}$ the formula

$$\bigvee_{1 \leq \ell_1 < \dots < \ell_k \leq m} (\varphi_{\ell_1} \wedge \dots \wedge \varphi_{\ell_k}).$$

If $(\psi_{1,} \dots, \psi_n)$ is another collection of formulas, we abbreviate by

$$(\varphi_{1,} \dots, \varphi_m) \leftrightarrow (\psi_{1,} \dots, \psi_n)$$

the formula

$$\bigwedge_{k=1}^{\max(m,n)} \varphi^{(k)} \leftrightarrow \psi^{(k)}.$$

The property (C1) is therefore expressed by

$$\begin{aligned} &((\varphi_{1,} \dots, \varphi_m) \leftrightarrow (\psi_{1,} \dots, \psi_n)) \rightarrow \\ &((\bigwedge_{i=1}^{m} L_{\alpha_i} \varphi_i) \bigwedge (\bigwedge_{j=1}^{n-1} M_{\beta_j} \psi_j) \rightarrow L_{(\alpha_1 + \dots + \alpha_m) - (\beta_1 + \dots \beta_{n-1})} \psi_n) \end{aligned} \tag{B}$$

We denote by $\mathcal{A}^+$ the system consisting of (A0), (A1), (A2), (A5), (A6) and (B). Our main result can now be stated:

[2]Here we adopt the convention that $E^{(k)} = \emptyset$ if $k > m$, and similarly $F^{(k)} = \emptyset$ for $k > n$.

Theorem: $\mathcal{A}^+$ *is a sound and complete axiomatization of* $\mathcal{M}$, *i.e.*

$$\vdash_{\mathcal{A}^+} \varphi \quad \Leftrightarrow \quad \mathcal{M} \models \varphi.$$

Before we elaborate on the theorem, we check that the schema (B) implies the schemata (A3), (A9), (A10) and (A11), or equivalently, that the property (C1) implies the properties (i)-(iv), and in fact more general versions of them.

It is clear that property (i) is a special case of (C1), by taking $m = 2$, $n = 1$, $E_1 = A$, $E_2 = B$ and $F_1 = A \cup B$. It is also not hard to show that (i) and (ii) are equivalent, and that (iii) and (iv) are equivalent. The remaining link is therefore:

Proposition 3: *Property (iii) follows from property (C1).*

The following proposition expresses some properties of the E_α operators.

Proposition 4: *From* $\mathcal{A}^+$ *the following axiom schemata can be derived:*

$$E_\alpha\varphi \leftrightarrow E_{1-\alpha}\neg\varphi \tag{A12}$$

$$E_\alpha\varphi \rightarrow \neg E_\beta\varphi \qquad \beta \neq \alpha \tag{A13}$$

Schema (A12) is an exact rendering of the complementation axiom, and (A13) says in effect that probability values are unique. Importantly, the *existence* of a probability value for each formula, as opposed to its *uniqueness*, cannot be expressed directly in our finitary language.

The property (C1), or its syntactical formulation (B), is very close in spirit to the sufficient condition of Kraft, Pratt and Seidenberg (1959) and Scott (1964) for a "more or equally probable than" relation on events $\succeq$ to be represented by a probability measure[3]. The condition says that if the sum of the characteristic functions of $E_1, \ldots, E_m$ equals that of $F_1, \ldots, F_m$ and

$$E_i \succeq F_i \quad i = 1, \ldots m-1$$

then $F_m \succeq E_m$. De Finetti had previously assumed that the simpler condition

$$E \succeq F \quad \Leftrightarrow \quad E \cup G \succeq F \cup G$$

[3]The relation $\succeq$ satisfies $\emptyset \not\succeq \Omega$ and $E \succeq \emptyset$ for every event $E \subseteq \Omega$, and every two events E and F are comparable – either $E \succeq F$ or $F \succeq E$.

for events G disjoint from both E and F, would suffice. However, a counterexample of Kraft, Pratt and Seidenberg (1959) showed that de Finetti's condition is not sufficient for the existence of a probability representation.

Qualitative probability relations have been investigated in some logic papers. Following Segerberg (1971), Gärdenfors (1975) introduced the binary relation $\succeq$ into a propositional language, and was thus able to translate the theory of qualitative probability relations, including Scott's condition, into syntactical terms. He states and proves completeness and decidability theorems for his system.

The property (C1) is also close in spirit to the necessary and sufficient condition found by Suppes and Zanotti (1989) for the existence of a probability μ between a pair (μ_*, μ^*) of lower and upper set functions $\mu_* \leq \mu^*$, of which the lower is super-additive and the upper sub-additive: For disjoint A and B,

$$\mu_*(A) + \mu_*(B) \leq \mu_*(A \cup B) \leq \mu^*(A \cup B) \leq \mu^*(A) + \mu^*(B).$$

The necessary and sufficient condition for the existence of a probability μ satisfying $\mu_* \leq \mu \leq \mu^*$ is that if the sum of the characteristic functions of $E_1, \ldots, E_m$ equals that of $F_1, \ldots, F_n$, then

$$\sum_{i=1}^{m} \mu_*(E_i) \leq \sum_{j=1}^{n} \mu^*(F_j)$$

Without this condition, counterexamples by Walley (1981) and Papamarcou and Fine (1986) show that a separating μ need not exist.

All these technical conditions share a common feature, i.e. they make it possible to employ some version of the separation theorem (or the Hahn-Banach theorem). In our case as well, the proof of the theorem will use a general version of the theorem of the alternative in convex analysis. This approach is coupled with the method of *filtration,* which has been used elsewhere in modal epistemic logic to prove the completeness of systems that are not necessarily *strongly* complete (e.g. Halpern and Moses' (1992) or Lismont and Mongin's (1994) common belief logics). With this technique, completeness is proved "formula by formula: one fixes the formula φ for which the implication

$$\models \varphi \quad \Rightarrow \quad \vdash \varphi$$

should hold, and proceeds to construct the finite space of maximally consistent sets of formulas in the sub-language $\mathcal{L}[\varphi]$ generated by φ, up to some finite depth. As it turns out, there is an *algorithm* to construct this space, which as a corollary delivers the *decidability* of the system.

The proof will be detailed in the full paper. There, we will also introduce the multi-agent extension of the system, as well as possible postulates for introspection in the probabilistic context.

3 References

Aumann, R.J. (1995), "Interactive Epistemology", *Discussion paper* no. 67, Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem.

Chellas, B. (1980), *Modal Logic,* Cambridge, Cambridge University Press.

Fagin, R. and J.Y. Halpern (1994), "Reasoning about Knowledge and Probability", *Journal of the Association for Computing Machinery* **41**, p. 340-367. Preliminary version in M.Y. Vardi (ed.), *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge* (TARK), Los Altos, Ca., Morgan Kaufmann, 1988, p. 277-293.

Fagin, R., J.Y. Halpern and N. Megiddo (1990), "A Logic for Reasoning about Probabilities", *Inf. Comput.* **87**, p. 78-128. Preliminary version in *Proceedings of the 1988 Symposium on Logic in Computer Science.*

Gaifman, H. (1964), "Concerning Measures in First Order Calculi", *Israel Journal of Mathematics,* **2**, p. 1-18.

Gärdenfors, P. (1975), "Qualitative Probability as an Intensional Logic", *Journal of Philosophical Logic* **4**, p. 171-185.

Halpern, J.Y. and Y. Moses (1992), "A Guide to Completeness and Complexity for Modal Logics of Knowledge and Beliefs", *Artificial Intelligence,* **54**, p. 319-379.

Harsanyi, J.C. (1967-68), "Games with Incomplete Information Played by Bayesian Players" parts I-III, *Management Science*, **14**, 159-182, 320-334, 486-502.

Kraft, C.H., J.W. Pratt and A. Seidenberg, "Intuitive Probability on Finite Sets", *Annals of Mathematical Statistics*, **30**, p. 408-419.

Lismont, L. and P. Mongin (1994), "A Non-Minimal but very Weak Axiomatization of Common Belief", *Artificial Intelligence*, **70**, p. 363-374.

Papamarcou, A. and T.L. Fine (1986), "A Note on Undominated Lower Probabilities", *The Annals of Probability*, **14**, p. 710-723.

Scott, D. (1964), "Measurement Structures and Linear Inequalities", *Journal of Mathematical Psychology*, **1**, p. 223-247.

Scott, D. and P. Krauss (1966), "Assigning Probabilities to Logical Formulas", in J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*, Amsterdam, North Holland, p. 219-264.

Segerberg, K. (1971), "Qualitative Probability in Modal Setting", *Proc. of the 2d Scan. Log. Symp.* (ed. Fenstad), Amsterdam.

Suppes, P. and M. Zanotti (1989), "Conditions on Upper and Lower Probabilities to Imply Probabilities", *Erkenntnis*, **31**, p. 323-345.

Walley, P. (1981), "Coherent Lower (and Upper) Probabilities", *Technical Report*, Department of Statistics, University of Warwick, England.