Int J Game Theory (1999) 28:301–314

International Journal of Game Theory

Interactive epistemology II: Probability*

Robert J. Aumann

Center for Rationality and Interactive Decision Theory, The Hebrew University, Feldman Building, Givat Ram, 91904 Jerusalem, Israel (e-mail: raumann@math.huji.ac.il)

Abstract. Formal Interactive Epistemology deals with the logic of knowledge and belief when there is more than one agent or "player." One is interested not only in each person's knowledge and beliefs about substantive matters, but also in his knowledge and beliefs about the others' knowledge and beliefs. This paper examines two parallel approaches to the subject. The first is the *semantic*, in which knowledge and beliefs are represented by a space Ω of *states of the world*, and for each player i, partitions $\mathscr{I}_i$ of Ω and probability distributions $\pi_i(\cdot\,;\omega)$ on Ω for each state ω of the world. The atom of $\mathscr{I}_i$ containing a given state ω represents i's knowledge at that state – the set of those other states that i cannot distinguish from ω; the probability distributions $\pi_i(\cdot\,;\omega)$ represents i's beliefs at the state ω. The second is the *syntactic* approach, in which beliefs are embodied in sentences constructed according to certain syntactic rules. This paper examines the relation between the two approaches, and shows that they are in a sense equivalent.

In game theory and economics, the semantic approach has heretofore been most prevalent. A question that often arises in this connection is whether, in what sense, and why the space Ω, the partitions $\mathscr{I}_i$, and the probability distributions $\pi_i(\cdot\,;\omega)$ can be taken as given and commonly known by the players. An answer to this question is provided by the syntactic approach.

* This paper is based on notes originally distributed in connection with a sequence of six lectures on Interactive Epistemology given at the Cowles Foundation for Research in Economics, Yale University, in the spring of 1989, and expanded and corrected several times since then. Previous versions were distributed in connection with the tutorial and workshop on "Knowledge and Game Theory" held at the Summer Institute in Game Theory, State University of New York at Stony Brook, July 20–24, 1992, and in connection with the Fifth Jerusalem Summer School in Economic Theory, on "Rationality of Action and Belief in the Economy," June 13–23, 1994. The author is grateful to Prof. Dov Samet for important help in preparing the current version. Research support from the USNSF under grant SBR-9730205 and previous grants is gratefully acknowledged.

Key words: Epistemology, interactive epistemology, probability, semantic, syntactic, model

11. Introduction

In interactive contexts like game theory and economics, it is important to consider what each player knows and believes about what the other players know and believe. Two different formalisms – the *semantic* and the *syntactic* – are available for this purpose. A companion paper (Aumann 1999, henceforth [A]) discusses and relates the two formalisms in the context of *knowledge*. Here we extend that analysis to include issues of belief – i.e., probability.

The semantic formalism consists of a "partition structure." For knowledge only [A], this consists of a space Ω of *states of the world* (or simply *states*), together with a partition of Ω for each player i. To deal with probability as well, one adds a probability distribution $\pi_i(\cdot;\omega)$ on Ω for each player i and each state ω. The atoms of i's partition are his *information sets*[1]; Ω is called the *universe*. Like in probability theory, *events* are subsets[2] of Ω; intuitively, an event E is identified with the set of all those states at which the event obtains. Thus E obtains at a state ω if and only if $\omega \in E$. Player i's probability for event E at state ω is represented by $\pi_i(E;\omega)$. If α is between 0 and 1 and E is an event, then "i's probability for E is at least α" is itself an event, denoted $P_i^\alpha E$; explicitly, it is the set of all states ω at which $\pi_i(E;\omega) \geq \alpha$. Similarly, "$i$ knows E" is an event, denoted K_iE; explicitly, ω is in K_iE if and only if the information set of i containing ω is included in E.

The syntactic formalism, on the other hand, is built on propositions, expressed in a formal language. The language has logical operators and connectives, operators k_i expressing knowledge, and operators p_i^α expressing beliefs. If e is a sentence, then k_ie and $p_i^\alpha e$ are also sentences; k_ie means "i knows, e," and $p_i^\alpha e$ means "i has probability at least α for e." The operators k_i and p_i^α can be iterated: The sentence $k_jp_i^\alpha e$ means "j knows that i's probability for j knowing e is at least α." Logical relations between the various propositions are expressed by formal rules.

There is a rough correspondence between the two formalisms: Events correspond to sentences, unions to disjunctions, intersections to conjunctions, inclusions to implications, complementation to negation, semantic knowledge operators K_i and belief operators P_i^α to syntactic knowledge operators k_i and belief operators p_i^α. But the correspondence really *is* quite rough; for example, only some – not all – events correspond to syntactically admissible sentences.

While the semantic formalism is the more convenient and widely used of the two, it is conceptually not quite straightforward. One question that often arises is, what do the players know about the formalism itself? Does each know the others' partitions, and does he know the others' probability measures as functions of ω? If so, from where does this knowledge derive?

[1] I.e., he can distinguish between states ω and ω' if and only if they are in different atoms of his partition.

[2] But unlike in [A], not *every* subset of Ω is an event; certain measurability conditions must be met. See Section 1.

For the case of knowledge, this question is answered by constructing a unique *canonical* semantic partition structure, in terms of the syntactic formalism; see [A]. The main purpose of this paper is to do the same in the expanded context that includes also probability.

The underlying idea is as in [A]. A state ω in the canonical semantic formalism is *defined* as a list of (syntactic) sentences that is complete and consistent in the appropriate senses; intuitively, the list is the set of all sentences that "hold at" ω. Also as in [A], the concept of a semantic *model* for a list of sentences plays a central role; a list is consistent if and only if it has a model.

But there are some significant divergences from [A]. For one thing, probability involves measurability; though this offers no particular difficulty, it must be dealt with. More important, in [A] the definition of "consistency" is syntactic, and one *proves* that a list of sentences is consistent if and only if it has a model. Here – in the context of probability – we did not succeed in formulating a satisfactory syntactic definition of consistency that would enable the proof of such a result. We therefore *define* a consistent list of sentences as one that has a model. This and other conceptual issues will be discussed in Sections 15 and 17.

While we have tried to make this paper self-contained, it does have important ties to [A], both conceptual and formal. To make it easier to refer to [A], we continue [A]'s numbering system here: The current paper starts with Section 11, so that a reference to 1.8, say, refers to [A]. For notations that are not explained here, the reader is referred to [A].

The remaining sections are numbered roughly as in [A], with the number 10 added. Thus Section 12 here describes the semantic knowledge-belief formalism, whereas Section 2 in [A] describes the (multi-player) semantic knowledge formalism. Sections 13 and 3 are conceptual discussions of the semantic formalisms. Sections 14 and 4 respectively set forth the syntactic knowledge-belief formalism, and the syntactic knowledge formalism; Sections 15 and 5 discuss these constructions. Sections 16 and 6 respectively construct the canonical semantic knowledge-belief and knowledge systems; again, Sections 17 and 7 discuss these constructions. Like Sections 8 and 9, Sections 18 and 19 explore and justify the foregoing material, using mathematical tools. Specifically, we show that there is a rough "isomorphism" between the two formalisms, similar to that established in Section 8 for the case of knowledge. Also, we show that the canonical system constructed in Section 16 is indeed a knowledge-belief system as defined in Section 12 (the corresponding statement for knowledge systems is immediate). The paper closes with a discussion in Section 20, roughly parallel to the discussion of knowledge formalisms in Section 10.

12. Semantic knowledge-belief formalisms

Recall that a *semantic knowledge system* (Section 2) consists of a *universe* Ω, a *population* N, and a *knowledge function* κ_i for each individual i in N. Define a *finite semantic knowledge-belief system* to consist of a semantic knowledge system with a finite universe, and, for each individual i and state ω, a probability measure $\pi_i(\cdot;\omega)$ on the field $\mathscr{E}$ of subsets of Ω (called *events*). The interpretation is that at state ω, individual i ascribes probability $\pi_i(E;\omega)$ to event E. Set

$$P_i^\alpha E := \{\omega : \pi_i(E;\omega) \geq \alpha\};\tag{12.1}$$

$P_i^\alpha E$ is the event that i ascribes probability at least α to E. Assume that

$$K_i E \subset P_i^1 E, \quad \text{and}\tag{12.2}$$

$$P_i^\alpha E \subset K_i P_i^\alpha E.\tag{12.3}$$

Intuitively, 12.2 says that if i knows something, then he assigns it probability 1; 12.3, that he knows the probabilities that he assigns to events. Technically, 12.2 says that $\pi_i(\cdot\,;\omega)$ is concentrated on i's information set $\mathbf{I}(\omega)$; 12.3, that $\pi_i(E;\cdot)$ is $\mathscr{K}_i$-measurable[3].

In the general (not necessarily finite) case, define a *knowledge-belief system* as a knowledge system $\{\Omega, N, \{\kappa_i\}_{i\in N}\}$, together with a sigma-field $\mathscr{F}_\sigma$ of sets in Ω (called *events*), and, for each individual i and state ω, a probability measure $\pi_i(\cdot\,;\omega)$ on $\mathscr{F}_\sigma$ such that

$$\text{the atoms of the } \mathscr{I}_i \text{ are } \mathscr{F}_\sigma\text{-measurable, and}\tag{12.4}$$

$$\pi_i(E;\omega) \text{ is } \mathscr{F}_\sigma\text{-measurable in } \omega \text{ for each fixed } E \text{ in } \mathscr{F}_\sigma.\tag{12.5}$$

As before, $\pi_i(E;\omega)$ signifies i's probability for E at ω; and we assume 12.2 and 12.3, where P_i^α is defined by 12.1.

For future reference, note that

$$\pi_i(E;\omega) = \sup\{\text{rational } \alpha : \omega \in P_i^\alpha E\}.\tag{12.6}$$

Indeed, fix ω, i and E. By 12.1, for any α (not necessarily rational) we have

$$\omega \in P_i^\alpha E \quad \text{iff } \pi_i(E;\omega) \geq \alpha.\tag{12.7}$$

Therefore $\pi_i(E;\omega)$ is $\geq$ any rational α for which $\omega \in P_i^\alpha E$, which establishes $\geq$ in 12.6. On the other hand, if a rational β is $\leq \pi_i(E;\omega)$, then by 12.7, $\omega \in P_i^\beta E$, so $\beta \leq \sup\{\alpha : \omega \in P_i^\alpha E\}$; this establishes $\leq$ in 12.6. ■

13. Discussion

First, note that as presented here, the beliefs of the players depend explicitly on their information; in effect, each player's probability is concentrated on his information set. This is in contrast to formulations (e.g., Aumann (1987)) in which the probability distributions are initially given on the entire state space Ω, and then the players compute posterior probabilities, conditional on their information. Proceeding as we do here emphasizes that we are analyzing a single moment of time, at which the players have the information that they have; we are interested in the beliefs of the players at that time, and at that time only.

[3] This means that if ω' and ω are in the same atom of the information partition $\mathscr{I}_i$, then $P_i(F,\omega') = P_i(F,\omega)$.

Next, the conceptual issues discussed in Section 3 arise here also. Basically, the problem is to justify the implicit assumption that the players know the model itself, including the partitions of the other players and their probability distributions (for each of their information sets). The problem could be resolved by presenting an explicit canonical semantic formalism, with explicit descriptions of the states, partitions, and probabilities. This, indeed, is what we do in Section 16.

14. The syntactic knowledge–belief formalism

Start with a finite population N, and a *keyboard* consisting of an alphabet $\mathfrak{X} := \{x, y, z, \ldots\}$ and the symbols $\vee, \neg,), (, k_i$, and p_i^α, where i ranges over the individuals in N and α over the rationals between 0 and 1 inclusive. A *formula* is defined as a finite string of symbols obtained by applying finitely often, in some order, the following rules:

$$\text{Every letter in the alphabet is a formula.} \tag{14.1}$$

$$\text{If } f \text{ and } g \text{ are formulas, so is } (f) \vee (g). \tag{14.2}$$

$$\text{If } f \text{ is a formula, so are } \neg(f), k_i(f), \text{ and } p_i^\alpha(f) \text{ for each } i \text{ and each } \alpha. \tag{14.3}$$

We often omit parentheses, and in particular write $p_i^\alpha f$ for $p_i^\alpha(f)$; intuitively, $p_i^\alpha f$ means "i ascribes probability at least α to f." Adding the probability operators p_i^α to the language enables us to refer to the players' beliefs, in addition to the elements treated in Section 4.

The set of all formulas with the given population N and alphabet $\mathfrak{X}$ is called a *syntax*, and is denoted $\mathfrak{S}(N, \mathfrak{X})$, or just $\mathfrak{S}$. Assume that N and $\mathfrak{X}$ are finite or denumerable; it follows that $\mathfrak{S}$ is denumerable. Define a *representation* of $\mathfrak{S}$ as a knowledge–belief system, $\hat{\Pi}$, together with a function $\varphi : \mathfrak{S} \to \hat{\mathscr{F}}_\sigma$ such that for all f, g, i, and α,

$$\varphi(\neg f) = \sim\varphi(f), \quad \varphi(f \vee g) = \varphi(f) \cup \varphi(g),$$
$$\varphi(k_i f) = \hat{K}_i\varphi(f), \quad \text{and} \quad \varphi(p_i^\alpha f) = \hat{P}_i^\alpha\varphi(f), \tag{14.4}$$

where $\hat{\mathscr{F}}_\sigma$ is the σ-field of events in $\hat{\Pi}$, and $\hat{K}_i$, $\hat{P}_i^\alpha$, are the corresponding semantic knowledge and belief operators. Given a representation $(\hat{\Pi}, \varphi)$ and a state $\hat{\omega}$ in $\hat{\Pi}$, a formula f is said to *hold at* $\hat{\omega}$ if $\hat{\omega}$ is in $\varphi(f)$. A *list* $\mathfrak{L}$ is a set of formulas. A *model* for $\mathfrak{L}$ is a representation φ and a state $\hat{\omega}$ in $\hat{\Pi}$ such that every formula in the list holds at $\hat{\omega}$. A formula f is a *consequence* of (or *follows* from) $\mathfrak{L}$ if every model for $\mathfrak{L}$ is also a model for $\{f\}$.

15. Discussion

In Section 5 we discussed and interpreted the syntax of pure knowledge, as set forth in Section 4; much of that discussion applies, mutatis mutandis, to the knowledge-belief syntax set forth in Section 14.

Each of these two syntaxes provides both a *grammar* and a *logic*. The grammar tells us how to construct "well-formed formulas," i.e., meaningful sentences; the logic, how to deduce formulas from one another.

The two grammars are entirely analogous. Both are embodied in the descriptions of the keyboards, and in the rules for constructing formulas from keyboard elements (14.1 through 14.3 for knowledge-belief, 4.1 through 4.3 for knowledge). Indeed, the only difference is that the knowledge-belief grammar has probability operators p_i^α in addition to the other elements.

But the two logics are quite different, in several ways. In Section 4 – which treats knowledge only – the fundamental notion of consequence is defined in purely syntactic terms: g is a *consequence* of f if it follows from f by repeated use of certain axioms and rules of deduction[4]. Moreover, it is finitary, in that the hypothesis comprises just one formula f; and though one *can* speak of g following from a list $\mathfrak{L}$ of formulas, that is tantamount to g following from a conjunction of *finitely* many formulas in $\mathfrak{L}$.

In contrast, in Section 14 the notion of consequence is infinitary: A formula can follow from a list $\mathfrak{L}$ without following from any finite sublist. For example, $p^{1/2}x$ follows from all the $p^\alpha x$ with $\alpha < 1/2$, but not from any finite number of them. Moreover, the definition is no longer purely syntactic; it involves "models," which are essentially semantic.

While these differences are conceptually significant, their practical effect is limited. Though infinitary deductions are in principle possible, we know of none in actual game-theoretic or economic applications of epistemological formalisms. As for the issue of syntax versus semantics in defining "consequence," practically speaking it shouldn't matter. In the case of knowledge, it indeed doesn't matter; by 9.4, the two approaches are equivalent. In the case of probability (knowledge-belief), we have not succeeded in developing a deductive logic that allows us to establish such a result formally. Conceptually, it would certainly be desirable to do so; and for a "purely" syntactic logic, without any semantic component, it is indispensable. But even without this, the definition of tautology in Section 14 does, practically speaking, provide a coherent logic for the syntactic grammar.

Section 16 uses the syntactic formalism of Section 14 to construct an explicit canonical semantic knowledge-belief system. But it should be remembered that the syntactic formalism is important not only as a tool for constructing the canonical semantic formalism, but also in its own right, as a formalism of deduction. In this formalism, there are no explicit states; one simply deduces true statements from other true statements. As noted in the introduction, this mode of deduction is in a sense more compelling – has more immediacy – than the semantic mode.

16. The canonical semantic knowledge-belief system

As in Section 14, assume given a finite population N and an alphabet $\mathfrak{X}$. Call a list $\mathfrak{L}$ of formulas *closed* if it contains all its consequences; *coherent*, if

$$\neg f \in \mathfrak{L} \quad \text{implies} \quad f \notin \mathfrak{L}; \tag{16.1}$$

[4] That is what "strong closure" amounts to.

and *complete* if

$$f \notin \mathfrak{L} \quad \text{implies} \quad \neg f \in \mathfrak{L}. \tag{16.2}$$

A *state* is a closed, coherent, and complete list. Denote the set of all states $\Pi(N, \mathfrak{X})$, or simply Π. For all individuals i, define a knowledge function κ_i on Π by specifying that for all states ω,

$$\kappa_i(\omega) \text{ is the set of all formulas in } \omega \text{ that start with } k_i. \tag{16.3}$$

For each formula f, define an event E_f (a subset of Π) by

$$E_f := \{\omega \in \Omega : f \in \omega\}. \tag{16.4}$$

Let $\mathscr{F}_\sigma$ be the σ-field generated[5] by the events E_f. For each f, i, and ω, define

$$\pi_i(E_f; \omega) := \sup(\alpha : p_i^\alpha f \in \omega). \tag{16.5}$$

We will show below (Section 18) that $\pi_i(\,\cdot\,;\omega)$ is well-defined on the events E_f for each ω and i, and extends uniquely to a σ-additive probability measure on $\mathscr{F}_\sigma$, which is also denoted $\pi_i(\,\cdot\,;\omega)$; moreover (Section 19), that the system comprising Π, the κ_i, the σ-field $\mathscr{F}_\sigma$, and the probability measures $\pi_i(\,\cdot\,;\omega)$ satisfies 12.2 through 12.5, and so is a knowledge-belief system. We call it the *canonical semantic knowledge-belief system* for $(N, \mathfrak{X})$, or simply the *canonical system*.

17. Discussion

As before, much of the discussion of the canonical semantic knowledge formalism applies, mutatis mutandis, also here; see Section 7.

We have seen (Section 15) that the logic of the syntactic knowledge-belief formalism depends on the notion of a semantic model, and so is not "purely" syntactic. This raises the question as to whether there is not some kind of circularity implicit in our construction of the canonical semantic formalism.

The answer is "no". Our aim, as stated at the end of Section 13, is to present an explicit canonical semantic formalism, with explicit descriptions of the states, partitions, and probabilities. The construction in Section 16 accomplishes this in a coherent and valid manner.

18. The canonical probabilities are well-defined and superadditive

Throughout this section and the next, ω denotes a state in the canonical system. In this section we prove that $\pi_i(\,\cdot\,;\omega)$ is well-defined on the events E_f (18.32), that it is finitely additive (18.4), that it extends uniquely to a sigma-additive measure on $\mathscr{F}_\sigma$ (18.56), and that it is a probability measure (18.6). In the process, we will begin, in the current knowledge-belief context, to establish

[5] The smallest sigma-field containing all the E_f.

a correspondence between syntax and semantics analogous to that established in Section 8 in the context of knowledge (18.1).

Let x be in the alphabet. Since ω is coherent, either $\neg x \notin \omega$ or $x \notin \omega$. Suppose w.l.o.g. that $x \notin \omega$. Since ω is closed, x is not a consequence of ω. Thus it is not true that every model for ω is also a model for x. So there is a model for ω that is not a model for x. In particular, there is a model for ω; denote it $(\hat{\Pi}, \varphi, \hat{\omega})$. Let $\hat{\pi}_i(\,\cdot\,;\hat{\omega})$ be the probabilities of i in $\hat{\Pi}$ at $\hat{\omega}$. Call an event of the form E_f *syntactic*, and denote the family of syntactic events by $\mathscr{F}$.

Proposition 18.1.

$$\sim E_f = E_{\neg f}, \tag{18.11}$$

$$E_f \cup E_g = E_{f \vee g}, \quad \text{and} \tag{18.12}$$

$$E_f \cap E_g = E_{f \wedge g}. \tag{18.13}$$

Proof: Analogous to that of 8.3, and so omitted.

Corollary 18.14. *The family* $\mathscr{F}$ *of syntactic events is a* field; *that is, it is closed under complementation, finite unions, and finite intersections.*

Lemma 18.2.

$$\textit{If } E_f \subset E_g, \textit{ then } \varphi(f) \subset \varphi(g), \tag{18.21}$$

$$f \in \omega \textit{ iff } \hat{\omega} \in \varphi(f), \quad \textit{and} \tag{18.22}$$

$$\hat{\pi}_i(\varphi(f);\hat{\omega}) = \sup\{\alpha : p_i^\alpha f \in \omega\}. \tag{18.23}$$

Proof: Suppose it is not the case that $\varphi(f) \subset \varphi(g)$; i.e., that there is an element $\hat{v}$ of $\varphi(f)\backslash\varphi(g)$. Let v be the list of all formulas that hold at $\hat{v}$. Then v is a consistent and complete list – i.e., a state in the canonical system Π. By construction, $f \in v$ and $g \notin v$, and the existence of such a state is incompatible with $E_f \subset E_g$. This establishes 18.21.

The "only if" part of 18.22 is what we mean by saying that $(\hat{\Pi}, \varphi, \hat{\omega})$ is a model for ω. For the "if" part, let $f \notin \omega$. Then $\neg f \in \omega$ by completeness (16.2). So by the "only if" part, $\hat{\omega} \in \varphi(\neg f) = \sim\varphi(f)$; that is, $\hat{\omega} \notin \varphi(f)$. This proves the contrapositive of the "if" part.

Finally, by 12.6 and 18.22, $\hat{\pi}_i(\varphi(f);\hat{\omega}) = \sup\{\alpha : \hat{\omega} \in \hat{P}_i^\alpha \varphi(f)\} = \sup\{\alpha : \hat{\omega} \in \varphi(p_i^\alpha f)\} = \sup\{\alpha : p_i^\alpha f \in \omega\}$, as asserted in 18.23. ■

Lemma 18.3. *If* $E_f \subset E_g$, *then* $\sup(\alpha : p_i^\alpha f \in \omega) \leq \sup(\alpha : p_i^\alpha g \in \omega)$ *for all* i *and* ω.

Proof: By 18.21, $\varphi(f) \subset \varphi(g)$, so $\hat{\pi}_i(\varphi(f),\hat{v}) \leq \hat{\pi}_i(\varphi(g),\hat{v})$; the lemma then follows from 18.23. ■

Corollary 18.31. *If* $E_f = E_g$, *then* $\sup(\alpha : p_i^\alpha f \in \omega) = \sup(\alpha : p_i^\alpha g \in \omega)$ *for all* i *and* ω.

Corollary 18.32. *The $\pi_i(\cdot\,;\omega)$, as defined by 16.5, are well-defined on the events E_f.*

Corollary 18.33. $\pi_i(E_f;\omega) = \hat{\pi}_i(\varphi(f);\hat{\omega})$.

Proof: Follows from 16.5 and 18.23. ■

Lemma 18.4. *The $\pi_i(\cdot\,;\omega)$ are finitely additive; that is, if $E_f \cap E_g = \varnothing$, then $\pi_i(E_f;\omega) + \pi_i(E_g;\omega) = \pi_i(E_f \cup E_g;\omega)$.*

Proof: $E_f \cap E_g = \varnothing$ and 18.1 yield $E_f \subset E_{\neg g}$. So by 18.21, $\varphi(f) \subset \varphi(\neg g) = {\sim}\varphi(g)$, so $\varphi(f) \cap \varphi(g) = \varnothing$, so by 18.12, 18.33, and 14.4, $\pi_i(E_f \cup E_g;\omega) = \pi_i(E_{f \vee g};\omega) = \hat{\pi}_i(\varphi(f \vee g);\hat{\omega}) = \hat{\pi}_i(\varphi(f) \cup \varphi(g);\hat{\omega}) = \hat{\pi}_i(\varphi(f);\hat{\omega}) + \hat{\pi}_i(\varphi(g);\hat{\omega}) = \pi_i(E_f;\omega) + \pi_i(E_g;\omega)$. ■

Lemma 18.5. *Let $g, f_1, f_2, \ldots$ be an infinite sequence of formulas such that*

$$E_{f_1} \subset E_{f_2} \subset \cdots \text{ and} \tag{18.51}$$

$$E_{f_1} \cup E_{f_2} \cup \cdots = E_g. \tag{18.52}$$

Then for each individual i and state ω,

$$\pi_i(E_g;\omega) = \lim_{n\to\infty} \pi_i(E_{f_n};\omega). \tag{18.53}$$

Proof: By 18.21,

$$\varphi(f_1) \subset \varphi(f_2) \subset \cdots \subset \varphi(g). \tag{18.54}$$

We claim that

$$\varphi(f_1) \cup \varphi(f_2) \cup \cdots = \varphi(g). \tag{18.55}$$

If not, then the last inclusion in 18.54 is strict; that is, there is a $\hat{v}$ in $\varphi(g)$ that is not in any of the $\varphi(f_n)$. Set $v := \{f : \hat{v} \in \varphi(f)\}$. Then v is a consistent and complete list of formulas – i.e., a state in Π. This state contains g but none of the f_n, contradicting 18.52; so 18.55 is proved. Then by 18.33, 18.54, 18.55, and the countable additivity of the $\hat{\pi}_i(\cdot\,;\hat{\omega})$,

$$\lim_{n\to\infty} \pi_i(E_{f_n};\omega) = \lim_{n\to\infty} \hat{\pi}_i(\varphi(f_n);\hat{\omega}) = \hat{\pi}_i(\varphi(g);\hat{\omega}) = \pi_i(E_g;\omega). \qquad ■$$

Corollary 18.56. *The definition (16.5) of $\pi_i(\cdot\,;\omega)$ extends uniquely to a σ-additive measure on $\mathscr{F}_\sigma$.*

Proof: By 18.14, the events E_f form a field. So, by Caratheodory's theorem, it suffices to show that $\pi_i(\cdot\,;\omega)$ is countably additive when restricted to events of the form E_f; i.e., that if $E_{g_1} \cup E_{g_2} \cdots = E_g$ and the E_{g_m} are mutually disjoint, then

$$\pi_i(E_{g_1};\omega) + \pi_i(E_{g_2};\omega) + \cdots = \pi_i(E_g;\omega).$$

To this end, set $f_n := g_1 \vee \cdots \vee g_n$, apply 18.5, 18.12, and 18.4, and conclude that

$$\pi_i(E_g;\omega) = \lim_{n\to\infty} \pi_i(E_{f_n};\omega) = \lim_{n\to\infty} \pi_i\left(\bigcup_{m=1}^{n} E_{g_m};\omega\right)$$

$$= \lim_{n\to\infty} \sum_{m=1}^{n} \pi_i(E_{g_m};\omega) = \sum_{m=1}^{\infty} \pi_i(E_{g_m};\omega). \qquad \blacksquare$$

Lemma 18.6. $\pi_i(\Pi;\omega) = 1$.

Proof: By completeness (16.2), each state contains x or $\neg x$, and so in any case $x \vee \neg x$. Therefore $\Pi = E_{x\vee\neg x}$, so by 11.9 and 13.23,

$$\pi_i(\Pi;\omega) = \pi_i(E_{x\vee\neg x};\omega) = \sup\{\alpha : p_i^\alpha(x \vee \neg x) \in \omega\} = \hat{\pi}_i(\varphi(x \vee \neg x);\hat{\omega})$$

$$= \hat{\pi}_i(\varphi(x) \cup \varphi(\neg x);\hat{\omega}) = \hat{\pi}_i(\varphi(x) \cup \varphi(\neg x);\hat{\omega})$$

$$= \hat{\pi}_i(\varphi(x) \cup \sim\varphi(x);\hat{\omega}) = \hat{\pi}_i(\hat{\Pi};\hat{\omega}) = 1. \qquad \blacksquare$$

19. The canonical system satisfies the conditions for a knowledge-belief system

In this section we show that the canonical system satisfies the conditions set forth in Section 4 for knowledge-belief systems: That an individual's probabilities are concentrated on his information set (19.5); that he knows his own probabilities (19.6); and that everything in sight is $\mathscr{F}_\sigma$-measurable (19.7 and 19.9). These four propositions establish Conditions (12.2) through (12.5) respectively.

We will also finish establishing, in the current knowledge-belief context, the correspondence between syntax and semantics (19.2 and 19.4).

The knowledge and probability operators of i in the canonical system will be denoted K_i and P_i^α (see 12.1). As in the previous section, ω denotes a state in the canonical system, $(\hat{\Pi}, \varphi, \hat{\omega})$ a model for ω, $\hat{\pi}_i(\,\cdot\,;\hat{\omega})$ the probabilities of i in $\hat{\Pi}$ at $\hat{\omega}$, $\hat{K}_i$ and $\hat{P}_i^\alpha$ the knowledge and probability operators in $\hat{\Pi}$.

We use 18.22 and 14.4 repeatedly in this section, without explicit mention.

Lemma 19.1.

$$\textit{If } k_i f \in \omega \textit{ then } f \in \omega, \tag{19.11}$$

$$\textit{if } \alpha \le \beta \textit{ and } p_i^\beta f \in \omega \textit{ then } p_i^\alpha f \in \omega, \textit{ and} \tag{19.12}$$

$$\textit{if } p_i^\alpha f \in \omega \textit{ then } k_i p_i^\alpha f \in \omega. \tag{19.13}$$

Proof: If $k_i f \in \omega$, then by 1.31, $\hat{\omega} \in \varphi(k_i f) = \hat{K}_i \varphi(f) \subset \varphi(f)$, so $f \in \omega$, proving 19.11. Next, let $p_i^\beta f \in \omega$, where $\alpha \le \beta$. Then $\hat{\omega} \in \varphi(p_i^\beta f) = \hat{P}_i^\beta \varphi(f) = \{\hat{v} : \hat{\pi}_i(\varphi(f);\hat{v}) \ge \beta\}$, so $\hat{\pi}_i(\varphi(f);\hat{\omega}) \ge \beta \ge \alpha$, so $\hat{\omega} \in \{\hat{v} : \hat{\pi}_i(\varphi(f);\hat{v}) \ge \alpha\} =$

$\hat{P}_i^\alpha\varphi(f) = \varphi(p_i^\alpha f)$, so $p_i^\alpha f \in \omega$, proving 19.12. Finally, if $p_i^\alpha f \in \omega$, then $\hat{\omega} \in \varphi(p_i^\alpha f) = \hat{P}_i^\alpha\varphi(f) \subset \hat{K}_i\hat{P}_i^\alpha\varphi(f) = \varphi(k_i p_i^\alpha f)$, by 12.3 applied to $\hat{\Pi}$; so $k_i p_i^\alpha f \in \omega$, proving 19.13. ■

Proposition 19.2. $K_i E_f = E_{k_i f}$.

Proof: We have $\omega \in K_i E_f$ iff $\mathbf{I}_i(\omega) \subset E_f$, and $\omega \in E_{k_i f}$ iff $k_i f \in \omega$. So we must show that $\mathbf{I}_i(\omega) \subset E_f$ iff $k_i f \in \omega$. If indeed $k_i f \in \omega$, then $k_i f$ is in each member of $\mathbf{I}_i(\omega)$, so f is in each member of $\mathbf{I}_i(\omega)$ (by 19.11, when applied to arbitrary states in $\mathbf{I}_i(\omega)$), so indeed $\mathbf{I}_i(\omega) \subset E_f$. For the opposite direction, suppose $k_i f \notin \omega$, so $\hat{\omega} \notin \varphi(k_i f) = \hat{K}_i\varphi(f)$. So $\hat{\mathbf{I}}_i(\hat{\omega})$ is not included in $\varphi(f)$. So there is an element $\hat{v}$ of $\hat{\Pi}$ with $\hat{\kappa}_i(\hat{v}) = \hat{\kappa}_i(\hat{\omega})$ and $\hat{v} \notin \varphi(f)$. Set

$$v := \{g : \hat{v} \in \varphi(g)\}; \tag{19.21}$$

in words, v is the list of all formulas that are true "at" $\hat{v}$. If $k_i h \in \omega$, then $\hat{\omega} \in \varphi(k_i h) = \hat{K}_i\varphi(h)$. Since $\hat{\kappa}_i(\hat{v}) = \hat{\kappa}_i(\hat{\omega})$, it follows that also $\hat{v} \in \hat{K}_i\varphi(h) = \varphi(k_i h)$, so $k_i h \in v$ by 19.11. If $k_i h \in v$, it may be seen similarly that $k_i h \in \omega$. So $\kappa_i(v) = \kappa_i(\omega)$. So $\mathbf{I}_i(v) = \mathbf{I}_i(\omega)$. But $\hat{v} \notin \varphi(f)$ and 19.21 yield $f \notin v$, so $v \notin E_f$. But by definition, $v \in \mathbf{I}_i(v) = \mathbf{I}_i(\omega)$. So it is not the case that $\mathbf{I}_i(\omega)$ is included in E_f. This completes the proof of the opposite direction. ■

Proposition 19.3. *If* $p_i^\beta f \in \omega$ *for all* $\beta < \alpha$, *then* $p_i^\alpha f \in \omega$.

Proof: By hypothesis, $\hat{\omega} \in \varphi(p_i^\beta f) = P_i^\beta\varphi(f)$ for all $\beta < \alpha$. So by 12.1, $\hat{\pi}_i(\varphi(f); \hat{\omega}) \geq \beta$ for all $\beta < \alpha$. So $\hat{\pi}_i(\varphi(f); \hat{\omega}) \geq \alpha$. So by 12.1, $\hat{\omega} \in P_i^\alpha\varphi(f) = \varphi(p_i^\alpha f)$. So $p_i^\alpha f \in \omega$. ■

Proposition 19.4. $P_i^\alpha E_f = E_{p_i^\alpha f}$ *for rational* α.

Proof: By 12.1 and 16.5, $\omega \in P_i^\alpha E_f$ iff $\alpha \leq \pi_i(E_f; \omega) := \sup(\beta : p_i^\beta f \in \omega)$; by 19.12 and 19.3, this happens iff $p_i^\alpha f \in \omega$, which in turn is iff $\omega \in E_{p_i^\alpha f}$. ■

Proposition 19.5. $K_i E \subset P_i^1 E$ *for all* E *in* $\mathscr{F}_\sigma$.

Proof: This says that $\pi_i(\cdot\,; \omega)$ is concentrated on $\mathbf{I}_i(\omega)$, i.e., that $\pi_i(\mathbf{I}_i(\omega); \omega) = 1$. Since $\mathbf{I}_i(\omega) = \bigcap_{k_i f \in \omega} E_{k_i f}$, and there are only denumerably many formulas in ω, it suffices to show that

$$\pi_i(E_{k_i f}; \omega) = 1 \text{ whenever } k_i f \in \omega.$$

Since $k_i f \in \omega$, 1.8 and 12.2 yield $\hat{\omega} \in \varphi(k_i f) = \hat{K}_i\varphi(f) = \hat{K}_i\hat{K}_i\varphi(f) = \hat{K}_i\varphi(k_i f) \subset \hat{P}_i^1\varphi(k_i f)$. So by 12.1, $\hat{\pi}_i(\varphi(k_i f); \hat{\omega}) = 1$. So by 18.33, $\pi_i(E_{k_i f}; \omega) = \hat{\pi}_i(\varphi(k_i f); \hat{\omega}) = 1$. ■

Proposition 19.6. $P_i^\alpha E \subset K_i P_i^\alpha E$ *for all* E *in* $\mathscr{F}_\sigma$.

Proof: By 19.13, 19.2 and 19.4, we know that $P_i^\alpha E \subset K_i P_i^\alpha E$ when E is syntactic (i.e., of the form E_f). Now the proposition says that

$$\text{if } \kappa_i(\omega) = \kappa_i(v), \quad \text{then } \pi_i(E;\omega) = \pi_i(E;v).$$

Fix ω and v such that $\kappa_i(\omega) = \kappa_i(v)$. We know that $\pi_i(E;\omega) = \pi_i(E;v)$ when E is syntactic. Since the syntactic events form a field, and this field generates $\mathscr{F}_\sigma$, it follows that any E in $\mathscr{F}_\sigma$ is approximable w.r.t. the sigma-additive measure $\pi_i(\cdot\,;\omega) + \pi_i(\cdot\,;v)$ by syntactic events; so $\pi_i(E;\omega) = \pi_i(E;v)$ for all E in $\mathscr{F}_\sigma$. ■

Proposition 19.7. *Then atoms of the $\mathscr{I}_i$ are $\mathscr{F}_\sigma$-measurable.*

Proof: Each atom of $\mathscr{I}_i$ is of the form $\mathbf{I}_i(\omega)$ for some ω, and $\mathbf{I}_i(\omega) = \bigcap_{k_i f \in \omega} E_{k_i f}$. ■

Call a sequence $E_1, E_2, \ldots$ of events *monotone* if $E_1 \subset E_2 \subset \cdots$ or $E_1 \supset E_2 \supset \cdots$; set $\lim_{n\to\infty} E_n := \bigcup_{n=1} E_n$ in the first case and $:= \bigcap_{n=1} E_n$ in the second. Call a family $\mathscr{G}$ *monotone* if $\lim_{n\to\infty} E_n \in \mathscr{G}$ whenever $E_1, E_2, \ldots$ is a monotone sequence in $\mathscr{G}$.

Lemma 19.8 *$\mathscr{F}_\sigma$ is the smallest monotone family that includes the field $\mathscr{F}$ of all the E_f.*

Proof: This holds for all σ-fields $\mathscr{F}_\sigma$ generated by a field $\mathscr{F}$ [Halmos 1950, p. 27, Theorem B]. ■

Proposition 19.9. *$\pi_i(E;\omega)$ is $\mathscr{F}_\sigma$-measurable in ω for each fixed E in $\mathscr{F}_\sigma$.*

Proof: We must show

$$\{\omega : \pi_i(E;\omega) \geq \alpha\} \in \mathscr{F}_\sigma \quad \text{for each rational } \alpha. \tag{19.91}$$

Now $\{\omega : \pi_i(E;\omega) \geq \alpha\} = P_i^\alpha E$. So when $E = E_f$, applying 19.4 yields $\{\omega : \pi_i(E;\omega) \geq \alpha\} = E_{p_i^\alpha f}$, which has the form E_g and so is indeed in $\mathscr{F}_\sigma$. Thus the family $\mathscr{G}$ of events satisfying 14.91 includes $\mathscr{F}$. It is therefore sufficient, by 19.8, to show that it is a monotone family. So let $E_1, E_2, \ldots$ be a monotone sequence in $\mathscr{G}$ with limit E. If the sequence is non-decreasing, then

$$\begin{aligned}\{\omega : \pi_i(E;\omega) \geq \alpha\} &= \{\omega : (\forall m)(\exists n)(\pi_i(E_n;\omega) \geq \alpha - (1/m)\} \\ &= \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \{\omega : \pi_i(E_n;\omega) \geq \alpha - (1/m)\} \in \mathscr{F}_\sigma,\end{aligned}$$

since $E_n \in \mathscr{G}$. If the sequence is non-increasing, then

$$\begin{aligned}\{\omega : \pi_i(E;\omega) \geq \alpha\} &= \{\omega : (\forall n)(\pi_i(E_n;\omega) \geq \alpha\} \\ &= \bigcap_{n=1}^{\infty} \{\omega : \pi_i(E_n;\omega) \geq \alpha\} \in \mathscr{F}_\sigma,\end{aligned}$$

since $E_n \in \mathscr{G}$. So $E \in \mathscr{G}$ in both cases, so $\mathscr{G}$ is indeed monotone. ■

20. Discussion

(a) The meaning of the knowledge operators K_i

As in the case of knowledge, the correct interpretation of the event K_iE is that "E follows logically from the syntactic events[6] that i knows." This follows from 8.51, which continues to hold – with much the same proof as before – in the context of this paper.

However, the analogue of 10.1 – that knowledge of an infinite disjunction is equivalent to knowledge of some finite subdisjunction – is *not* correct here. For example,

$$K_i(E_{p_j^{1/2}x} \cup E_{p_j^{1/3}x} \cup E_{p_j^{1/4}x} \cup \cdots) \neq K_i(E_{p_j^{1/2}x}) \cup K_i(E_{p_j^{1/3}x}) \cup K_i(E_{p_j^{1/4}x}) \cup \cdots;$$

the left side equals $K_i(E_{\neg p_j^0 x})$ – i.e., it says that i knows that j's probability for x is positive – whereas the right side says that for some α that is positive, i knows that j's probability for x is at least α – a much stronger statement. Unlike that of 8.51, the proof of 10.1 involves compactness arguments, which, because of the infinitary nature of the logic, do not apply here.

(b) An alternative interpretation of the letters of the alphabet

The discussion at 10(b) continues to apply here, with little or no change.

(c) Knowledge-belief hierarchies

In 10(c) we saw that the hierarchy approach to knowledge, though more cumbersome and complicated than the syntactic approach, amounts to the same thing. The situation is similar with regard to knowledge-belief hierarchies.

Historically, the hierarchy approach originated with probability systems; see Armbruster and Böge (1979), Böge and Eisele (1979) and Mertens and Zamir (1985). These systems are pure probability systems; knowledge does not enter. For simplicity, we confine attention to the case of two players. Like at 10(c), we start with a set $\mathfrak{H}$ of mutually exclusive and exhaustive "states of nature," which describe some aspect of reality (like tomorrow's temperature in Jerusalem) in terms not involving probability or knowledge. Then the hierarchy of a player i consists of the beliefs (probabilities) of i about the states of nature, i's probability distribution over the beliefs of the other player j about the states of nature, i's probability distribution over j's probability distributions over i's beliefs about the state of nature, and so on. Certain consistency conditions must be met. Moreover, j's probability distribution over i's beliefs about the states of nature may well be continuous; thus the next level of the hierarchy consists of a probability distribution over probability distributions, which requires selecting a topology on the space of probability distributions. So even without knowledge, matters quickly get rather complicated.

[6] Events of the form E_f.

But knowledge *is* an intrinsic part of the picture, and should not be ignored. For example, a player necessarily *knows* his own probabilities, in the sense of absolute knowledge, not only in the sense of probability 1. The appropriate object to examine is a *knowledge-belief* hierarchy. In the case of two players, a pair of hierarchies (one for each player) looks roughly as follows: At the first level, each player i has a knowledge-belief profile over the set $\mathfrak{H}$ of states of nature; that is, he knows that the true state of nature is in a certain subset h_i of $\mathfrak{H}$, and he has a probability distribution on h_i. The knowledge of the two players must be *consistent*; one player cannot know something that the other knows to be false. At the second level of the hierarchy, each player has some knowledge and beliefs about pairs consisting of elements of $\mathfrak{H}$ and the other player's first-level knowledge and beliefs; these second-level knowledge-belief profiles of the two players must be consistent with each other in a sense like that described for the first level, and each player's second stage profile must also be consistent with his first level profile (e.g., the first level must be the marginal of the second level when projected onto the first level). And so on, ad infinitum. A more precise description, in the spirit of 10(c), would of course be far more complicated.

All this is accomplished much more briefly and elegantly by the syntactic approach. If ω is a state of the world in the canonical knowledge-belief system Π, then $\kappa_i(\omega)$ contains *precisely* the same information as i's hierarchy in the hierarchy approach. Just list all the formulas that i knows to be true – which of course include all his probability assessments, all his probability assessments about others' probability assessments, and so on. No complicated consistency conditions, no topologies. Just a list of the formulas he knows, in no particular order.

As at 10(c), note that two states of the world ω and ω' are in the same element of i's partition of Π if and only if they correspond to the same knowledge-belief hierarchy of i. Thus i's knowledge-belief hierarchies correspond precisely to the atoms of his information partition; each can be read off from the other. In particular, it follows that the knowledge-belief hierarchy of any *one* player determines precisely the common knowledge component of the true state of the world.

References

Armbruster W, Böge W (1979) Bayesian game theory. In: Möschlin O, Pallaschke D (eds.) Game theory and related topics, North-Holland, Amsterdam

Aumann RJ (1987) Correlated equilibrium as an expression of Bayesian rationality. Econometrica 55:1–18

Aumann RJ (1999) Interactive epistemology I: Knowledge. International Journal of Game Theory 28: 263–300 (denoted [A] in this paper)

Böge W, Eisele T (1979) On solutions of Bayesian games. International Journal of Game Theory 8:193–215

Halmos PR (1950) Measure Theory. Van Nostrand, Princeton

Mertens JF, Zamir S (1985) Formulation of Bayesian analysis for games with incomplete information. International Journal of Game Theory 14:1–29