INFINITARY EPISTEMIC LOGIC

Aviad Heifetz

School of Mathematical Sciences
Tel Aviv University
Tel Aviv, Israel
e_mail: aviad@math.tau.ac.il

December 1993

ABSTRACT

It is known that a theory in S5-epistemic logic with several agents may have numerous models. This is because each such model specifies also what an agent knows about infinite intersections of events, while the expressive power of the logic is limited to finite conjunctions of formulas. We show that this asymmetry between syntax and semantics persists also when infinite conjunctions (up to some given cardinality) are permitted in the language. We develop a strengthened S5-axiomatic system for such infinitary logics, and prove a strong completeness theorem for them. Then we show that in every such logic there is always a theory with more than one model.

I am grateful to Dov Samet for fruitful discussions.

§1.Introduction

It is well established that models for the ordinary propositional calculus are essentially unique. A model is a set of states, where each formula in the language is associated with the set of states where it holds, such that conjunction of formulas corresponds to union of events and negation to complements. The model is determined once the truth or falsity of every atomic formula is fixed in each state.

The situation is different in epistemic logic when several agents are involved. Here, the language is enriched with a knowledge operator k^i for each agent i. For the S5 axiom system, the standard models are partition spaces (Aumann 1976, or, equivalently, Kripke structures where the possibility relation for each agent is an equivalence relation). The interpretation is that in a given state an agent considers as possible all those states in the member of his partition that contains the current state. This means that the formula $k^i\varphi$ holds true wherever the member of i's partition is contained in the event that corresponds to φ.

Partition spaces are used in game theory and economics with the implicit assumption that there is no uncertainty of one agent regarding the partitions of the others. Aumann (1989) justifies this assumption by showing, that the states of partition spaces can be constructed out of maximally consistent sets of formulas in epistemic logic. Thus, he claims, the relevant aspects of knowledge or lack of knowledge of every agent can be specified endogenously in each state.

There is one important aspect, however, in which this epistemic logic falls short of characterizing partition spaces: There are maximally consistent sets of formulas with more than one model. By this we mean that there may be two partition spaces, non of which can be embedded in the other (*), while both of which contain a point where exactly those formulas in the maximally consistent set obtain.

The reason for this phenomenon is that the partition of an agent determines his knowledge concerning any event, while generally not every event corresponds to some formula. For instance, if $\varphi_1, \varphi_2, \ldots$ are associated with $E_1, E_2, \ldots$, respectively, there may be no formula associated with $\bigcap_{n=1}^{\infty} E_n$, because there is no formula in the language which is the conjunction of all the φ_n-s. Furthermore, if an agent does not exclude the possibility of $E_1, E_2, \ldots$, sometimes it might be equally reasonable to assume that he excludes their intersection, as well as to assume that he does not. These two options would then be reflected by two different models for $\varphi_1, \varphi_2, \ldots$. Such an example (with three agents) was given by Fagin, Halpern and Vardi (1991). Another example, which will be useful in the sequel, appears in Heifetz and Samet (1993).

The conclusion is that partition spaces are more decisive and expressive than the logic that they model. Can this asymmetry be remedied by considering a more powerful logic? The flaw seems to lie in the lack of infinite conjunctions and disjunctions in the usual epistemic logic. We pursue to strengthen the logic exactly in this respect. We merge the axiom system and provability notions for infinitary propositional calculus, initiated by Karp (1964), with strengthened S5-axioms, and check the relations between

(*) and non of which is redundant – in different points of each space different sets of formulas are true.

the resulting logic and its semantics in partition spaces.

First, we present a strong completeness result that relies on Karp's (1964) distributive axiom, which is essentially a logical representation of the axiom of choice. Then, we present two major results – a positive one and a negative one. The positive result says that for every partition space there is an infinitary epistemic logic which is equally potent – for each state in the space, the set of formulas that obtain in this state has a unique model which is minimal and non-redundant. On the other hand, the negative result says that no infinitary language suffices to characterize every partition space: Once you fix the cardinality of sets of formulas that admit conjunctions, there will always be maximally consistent sets of formulas with more than one model. This means that the inadequacy of epistemic logic – either finitary or infinitary – in characterizing its models *is essential.*

Our theorems on the relations between syntax and semantics rely on the results already obtained for the semantical part. Fagin, Halpern and Vardi (1991) researched a hierarchical version of Aumann's logic (1989), augmented by certain infinite conjunctions – those representing common knowledge of formulas. Fagin, Geanakoplos, Halpern and Vardi (1992) and Heifetz (1992) introduced the simpler to handle set-theoretic building blocks rather than the logical ones in the hierarchical construction of states. The general case was studied by Fagin (1993) and by Heifetz and Samet (1993). The former investigated the possible complexity of Kripke structures and their size. The latter built universal spaces in which every partition space can be embedded. We will be using these universal spaces and their properties for proving our results.

The paper is organized as follows: in section §2 we present formally the syntax and semantics we will be using, and state the strong completeness result. In section §3 we weigh the relative expressive power of the logic versus its models. We prove the inadequacy of the infinitary epistemic logic in characterizing its models using a constructive example from Heifetz and Samet (1993). In section §4 we bring the results on universal partition spaces needed for the remaining proofs, which are detailed in section §5.

§2. Language, Axioms, Models

Throughout the analysis we fix the set I of agents and the set A of atomic formulas. There are at least two agents. We denote by $|X|$ the cardinality of the set X.

Let c be an infinite cardinal. The language $\mathcal{L}_c$ is the minimal set of formulas that contains A and is closed under the following rules:

1) If φ, ψ are formulas in $\mathcal{L}_c$ then so are $\neg\varphi$ and $k^i\varphi$ for every agent $i \in I$.
2) If $\{\varphi_\ell\}$ is a set of formulas of cardinality less than c then $\bigwedge \varphi_\ell$ is also a formula in $\mathcal{L}_c$.

As usual, $\bigvee \varphi_\ell$ will stand for $\neg(\bigwedge \neg\varphi_\ell)$, and $\varphi \to \psi$ will stand for $\neg\varphi \bigvee \psi$.

We will associate ordinal ranks to formulas, denoted $\text{rank}(\varphi)$, by the following rules:

1) $\text{rank}(\varphi) = 0$ for atomic φ.
2) $\text{rank}(k^i\varphi) = \text{rank}(\varphi) + 1$.
3) $\text{rank}(\neg\varphi) = \text{rank}(\varphi)$.
4) $\text{rank}(\bigwedge \varphi_\ell)$ is the supremum of the ranks of the φ_ℓ-s.

Whenever Γ is a set of formulas, we will denote by rank(Γ) the supremum of the ranks of the formulas in Γ.

The axiom schemes for $\mathcal{L}_c$ are the following:

$$(A1) \qquad \varphi \rightarrow (\psi \rightarrow \varphi)$$

$$(A2) \qquad (\varphi \rightarrow (\psi \rightarrow \rho)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \rho))$$

$$(A3) \qquad (\neg\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \varphi)$$

$$(A4) \qquad \bigwedge(\psi \rightarrow \varphi_\ell) \rightarrow (\psi \rightarrow \bigwedge \varphi_\ell)$$

$$(A5) \qquad \bigwedge \varphi_\ell \rightarrow \varphi_m \qquad \forall \varphi_m \in \{\varphi_\ell\}$$

$$(A6) \qquad \bigwedge_{\ell<d}\Big(\bigvee_{m<d} \varphi_{\ell m}\Big) \rightarrow \bigvee_{g\in d^d}\Big(\bigwedge_{m<d} \varphi_{\ell g(\ell)}\Big) \qquad \text{whenever } 2^d < c$$

$$(K1) \qquad k^i\varphi \rightarrow \varphi$$

$$(K2) \qquad k^i(\bigwedge \varphi_\ell \rightarrow \psi) \rightarrow (\bigwedge k^i\varphi_\ell \rightarrow k^i\psi)$$

$$(K3) \qquad k^i\varphi \rightarrow k^i k^i \varphi$$

$$(K4) \qquad \neg k^i\varphi \rightarrow k^i \neg k^i \varphi$$

Axiom schemes $(A1) - (A5)$ are the direct generalizations of the axioms for the usual propositional calculus, taken from Karp (1964). Axiom scheme $(A6)$ is Karp's distributive law expressing the axiom of choice ("if for every $\ell < d$ there is $m < d$ such that $\varphi_{\ell m}$ holds, then there is a choice function g from d to d such that all the $\varphi_{\ell g(\ell)}$-s hold at once"). Axiom schemes $(K1) - (K4)$ reinforce the usual $S5$ axioms for knowledge. (($K3$) follows from $(K1)$, $(K2)$ and $(K4)$ - see, for instance, Chellas (1980) or Hughes and Cresswell (1984). We bring $(K3)$ to ease the symmetry of the discussion in the sequel).

We will also use the following inference rules, adopted from Karp(1964):

Modus Ponens. From φ and $\varphi \to \psi$ infer ψ.

Conjunction. From $\{\varphi_\ell\}$ infer $\bigwedge \varphi_\ell$.

The set of *tautologies* is the minimal set of formulas that contains all the axioms of the form $(A1)-(A6)$, $(K1)-(K4)$ and is closed under Modus Ponens, Conjunction and the inference rule of

Logical Omniscience. From φ infer $k^i\varphi$.

Let Γ be a set of formulas in $\mathcal{L}_c$. A *proof* of φ from Γ is a sequence of formulas whose length is smaller than c and whose last formula is φ, such that each formula in the proof is either in Γ, a tautology, or inferred from previous formulas by Modus Ponens or Conjunction.

Whenever there is a proof of φ from Γ we write $\Gamma \vdash \varphi$. Γ is said to be consistent if one can not prove from Γ a formula and its negation. It is said to be maximally consistent if for every $\varphi \in \mathcal{L}_c$ it contains either φ or $\neg\varphi$.

A *model* for $\mathcal{L}_c$ is a space Ω with partitions Π^i for the agents, together with a mapping f from formulas in $\mathcal{L}_c$ to events (-subsets) of Ω, such that $\forall\varphi \in \mathcal{L}_c$

$$f(\neg\varphi) = f(\varphi)^c, \qquad f(\bigwedge \varphi_\ell) = \bigcap f(\varphi_\ell)$$

and $\forall i \in I$

$$f(k^i\varphi) = K^i(f(\varphi))$$

Where K^i is the knowledge operator on events in Ω

$$K^i(E) = \{\omega \in \Omega : \Pi^i(\omega) \subseteq E\}.$$

In particular, we say that Ω is a model of a set of formulas $\Gamma \subseteq \mathcal{L}_c$ if there is an $\omega \in \Omega$ where all the formulas of Γ obtain under the mapping f. We then say that Γ holds in ω.

Clearly, once f is defined for the atomic formulas, the partitions Π^i determine how f should act on every formula in $\mathcal{L}_{c'}$ for every cardinality c' – this can be easily verified by induction on the structure of formulas. Therefore, from now on we will specify a model by the triple $(\Omega,\ (\Pi^i)_{i\in I},\ f : A \to 2^\Omega)$, understanding that for a given $\mathcal{L}_c$, f is extended in the above way to all the formulas of $\mathcal{L}_c$.

A model is called *non-redundant* if any two points in it are separated by some formula in $\mathcal{L}_{c'}$ for some cardinality c' under the mapping f. We say that the model is *minimal* if there is no $\Omega' \subsetneq \Omega$ such that $\Pi^i(\omega') \subseteq \Omega'$ for all $\omega' \in \Omega'$, $i \in I$.

A model $(\Omega,\ (\Pi^i)_{i\in I},\ f : A \to 2^\Omega)$ is *epimorphic* to another model $(\overline{\Omega},\ (\overline{\Pi}^i)_{i\in I},\ \overline{f} : A \to 2^{\overline{\Omega}})$ if there is a map H of Ω onto $\overline{\Omega}$ such that $H(f(\varphi)) = \overline{f}(\varphi)$ for every atomic $\varphi \in A$, and such that for every $i \in I$ and $\omega \in \Omega$ we have $H(\Pi^i(\omega)) = \overline{\Pi}^i(H(\omega))$. Clearly, an epimorphism preserves the truth of formulas that hold in states. H is an *isomorphism* between the models whenever it is one-to-one.

We say that φ follows semantically from Γ – $\Gamma \models \varphi$ – if in any model Ω of $\mathcal{L}_c$, φ holds in $\omega \in \Omega$ whenever all the formulas of Γ hold in ω. It is straightforward to check that in such a case if $\Gamma \vdash \varphi$ then φ also holds in ω, and therefore that only consistent Γ may have models. This means that provability in $\mathcal{L}_c$ is strongly sound: $\Gamma \vdash \varphi$ implies $\Gamma \models \varphi$. As for strong completeness, we have the following result:

Theorem 2.1. In $\mathcal{L}_c$

$$\Gamma \models \varphi \quad \text{implies} \quad \Gamma \vdash \varphi$$

whenever c is greater than some lower bound that depends on rank$(\Gamma \cup \{\varphi\})$.

§3. Syntax versus Semantics

We are now ready to introduce our results regarding the relative expressive power of infinitary epistemic logic with respect to partition spaces. The first, positive result suggests that we can always find a logical setting which is as expressive as a given partition space.

Theorem 3.1. For every model $(\Omega,\ (\Pi^i)_{i\in I},\ f : A \to 2^{\Omega})$ there is a cardinal c, such that for every $\omega \in \Omega$ the set of formulas of $\mathcal{L}_c$ that obtain in ω (under the unique extension of f to $\mathcal{L}_c$) has a unique minimal, non-redundant model.

The second, negative result says that no $\mathcal{L}_c$ suffices to characterize all its possible models.

Theorem 3.2. For every infinite cardinal c, there is a maximally consistent set of formulas in $\mathcal{L}_c$ that has more than one minimal and non-redundant model.

Proof of Theorem 3.2. We use the following example from Heifetz and Samet (1993). Suppose there is a unique atomic formula φ in $\mathcal{L}_c$ and only two agents. (It will be transparent from what follows that the argument also holds when there are other atomic formulas and agents.) Consider the following model for $\mathcal{L}_c$. The space Ω will consist of triples (r, t^1, t^2). r is either φ or $\neg\varphi$, and t^1, t^2 are the types of agent 1 and 2, respectively. a type of an agent is a sequence of S-s and D-s of length $c+1$ (*). S stands for Sober, D for Drunk. Our agents are those hopelessly addicted to alcohol: Whenever they try to avoid it they finally fail and give up. Formally, for every type and every limit ordinal $\lambda \leq c$ there exists a point $\beta < \lambda$ such that there are only D-s in the sequence from β up to (not including) λ. The type is called λ-even if this β is even (**) and λ-odd otherwise.

Define the partition member of each agent in a point (r, t^1, t^2) such that the following properties obtain:

1) Each agent knows his own type. However:
2) He perceives the state of nature if and only if his type starts by S (=Sober);

(*) i.e., a type is a point in $\{S, D\}^{c+1}$, where $c+1$ is the set of all ordinals smaller or equal to c, when c is identified with the least ordinal with cardinality c.
(**) i.e., in the unique representation $\beta = \lambda' + n$, where λ' is a limit ordinal (or 0) and $n \in \mathrm{N}$, n is even.

3) He perceives whether the other agent is Sober or Drunk in the γ-coordinate if and only if he himself is Sober in the $\gamma+1$-coordinate;
4) He perceives whether the other agent is λ-even or λ-odd if and only if he is Sober in the λ-coordinate;
5) (applies only when α is not a limit ordinal:) In any case he does not perceive whether the other agent is Sober or Drunk in the last coordinate.

The crucial point in this example is that no combination of S-s and D-s up to (not including) a limit ordinal λ ever enables an agent to perceive whether the other agent is λ-even or λ-odd. This is because in such a combination he is always Drunk from some β on, so he can not exclude the possibility that the other agent stood the temptation longer than he did, and fell Drunk (up to λ) only at some later stage $\beta' > \beta$, where β' may be even as well as odd. Therefore, becoming Sober in stage λ enables the agent to exclude some types of the other agent that he can not exclude when he is Drunk in stage λ, so these two options are epistemically different. This epistemic difference applies also for the last level $\lambda = c$, but it can no longer be expressed by any formula in $\mathcal{L}_c$, because we would need a disjunction of c formulas to represent it.

The formal argument goes as follows. We make Ω (with the partitions described above) into a model by mapping φ to the set of all states where $r = \varphi$. This map f extends uniquely to all the formulas of $\mathcal{L}_{c'}$ for every cardinal c', and in particular to $\mathcal{L}_c$. Define the formulas s^i_γ (read: "agent i is sober in level γ") and e^i_λ (read: "agent i is λ-even) by transfinite induction in the following way:

$$s^i_0 = k^i\varphi \vee k^i\neg\varphi$$

$$s^i_{\gamma+1} = k^i s^j_\gamma \vee k^i \neg s^j_\gamma \qquad j \neq i$$

$$e^i_\lambda = \bigvee_{\substack{\beta \text{ even} \\ \beta<\lambda}} \Big(\bigwedge_{\beta\leq\gamma<\lambda} \neg s^i_\gamma \Big) \qquad \lambda \text{ a limit ordinal}$$

$$s^i_\lambda = k^i e^j_\lambda \vee k^i \neg e^j_\lambda \qquad j \neq i,\ \lambda \text{ a limit ordinal}$$

It is not hard to check that for all $\gamma < c$, s^i_γ holds under f exactly in those states where agent i is Sober in his γ-coordinte (the formal proof may be found in Heifetz and Samet (1993)). Furthermore, the formulas φ, s^i_γ or its negations that hold in a given state either prove or disprove also every other formula of $\mathcal{L}_c$ (this can be easily shown by induction on the structure of formulas in $\mathcal{L}_c$). Notice, however, that for $\beta = c$, e^i_c and s^i_c are no longer formulas in $\mathcal{L}_c$ but only in $\mathcal{L}_{c^+}$ (where c^+ is the successor cardinal to c), because in these formulas we have conjunctions and disjunctions over c sub-formulas. This means that any two states of Ω are separated by some formula in $\mathcal{L}_{c^+}$, so the model is non-redundant. Notice also that Ω is minimal – it has no proper sub-models. Now, consider any two states ω, ω' in Ω which differ only by their last coordinate of level c for agent i. Then the same maximally consistent set of formulas Γ in $\mathcal{L}_c$ holds in both of them. But s^i_c, which is outside $\mathcal{L}_c$, holds in one of these states and not in the other, so Ω is not epimorphic to itself under a map that maps ω to ω'. This means that Ω with ω and Ω with ω' are two non-isomorphic, minimal and non-redundant models for Γ. ∎

Remark 3.3. At first sight it might seem, that the above result depends on the fact that we always limit the allowed conjunctions by some cardinality. Notice, however, that given any set of formulas there is always a bound on the cardinality of conjunctions in the set. There may be no such bound only if we consider collections of formulas which are proper classes. But in such a case we might need to consider models which are also proper classes (or even "larger" objects), and then the example in the proof could be extended to exhibit a consistent proper class of formulas with more than one model. (see also the discussion in section §5 of Heifetz and Samet (1993).)

§4. Universal Partition Spaces

In this section we bring the results from Heifetz and Samet (1993) that we need in order to prove the remaining theorems. Parallels to some of these results, in different terms and setup, appear also in Fagin (1993).

Let $S = 2^A$ be the collection of truth assignments to the atomic formulas. Define, by transfinite induction, the following spaces:

$$W_0 = S$$

$$W_\alpha = \Big\{ (s,\ (t^j_\beta)_{\substack{\beta<\alpha \\ j\in I}}) \in S \times \prod_{\beta<\alpha} \left(2^{W_\beta}\right)^I : \ \forall \beta < \alpha \quad \forall i \in I$$

$$\text{(I)} \quad (s,\ (t^j_\gamma)_{\substack{\gamma<\beta \\ j\in I}}) \in t^i_\beta$$

$$\text{(II)} \quad (\bar{s},\ (\bar{t}^j_\gamma)_{\substack{\gamma<\beta \\ j\in I}}) \in t^i_\beta \Rightarrow \bar{t}^i_\gamma = t^i_\gamma \quad \forall \gamma < \beta$$

$$\text{(III)} \quad \forall \gamma < \beta \text{ the projection of } t^i_\beta \text{ on } W_\gamma \text{ is } t^i_\gamma \quad \Big\}$$

The partition Π^i_α of agent i on W_α is defined for every $w_\alpha = (s,\ (t^j_\beta)_{\substack{\beta<\alpha \\ j\in I}}) \in W_\alpha$ by

$$\Pi^i_\alpha(w_\alpha) = \{ (\bar{s},\ (\bar{t}^j_\beta)_{\substack{\beta<\alpha \\ j\in I}}) \in W_\alpha : \bar{t}^i_\beta = t^i_\beta \quad \forall \beta < \alpha \}.$$

Thus, condition (I) in the definition of W_α says that i never excludes the prevailing state, condition (II) says he knows his own beliefs, and condition (III) says that his different levels of beliefs are coherent with one another. Intuitively speaking, under the partition Π^i_α agent i does not know anything that is not explicitly specified by his characteristics in the given state. These characteristics always refer to his beliefs on the layers before α, but not on layer α itself.

Clearly, every W_α is a model for any $\mathcal{L}_c$ with the partitions Π^i_α, when atomic formulas are associated with events by the natural map

$$f_\alpha(\varphi) = \{ (s,\ (t^j_\beta)_{\substack{j\in I \\ \beta<\alpha}}) \in W_\alpha : \varphi \in s \},$$

and then f_α is uniquely extended to all of $\mathcal{L}_c$.

$\widehat{W}$ is said to be a subspace of W_α if for every $\widehat{w} \in \widehat{W}$ we have $\Pi^i_\alpha(\widehat{w}) \subseteq \widehat{W}$ for all $i \in I$, and furthermore $(\widehat{w}, (\Pi^i_\alpha(\widehat{w}))_{i \in I})$ is the unique extension of $\widehat{w}$ to a point in $\widehat{W}_{\alpha+1}$.

For any model $(\Omega, (\Pi^i)_{i\in I}, f : A \to 2^A)$ for $\mathcal{L}_c$ define now the maps H_α of the model into the W_α-s, by transfinite induction:

$$H_0(\omega) = \{\varphi \in A : \omega \in f(\varphi)\}$$

(Recall that $W_0 = 2^A$),

$$H_\alpha(\omega) = (H_0(\omega),\ (H_\beta(\Pi^i(\omega)))_{\substack{\beta<\alpha \\ i\in I}})$$

The following theorem is a corollary of the main result in Heifetz and Samet (1993).

Theorem 4.1. For every model $(\Omega, (\Pi^i)_{i\in I}, f : A \to 2^A)$ for $\mathcal{L}_c$ there is a (limit) ordinal λ, such that for all $\alpha \geq \lambda$, H_α is an epimorphism of the model onto a subspace of W_α. H_α is one-to-one if and only if the model is non-redundant.

§5. Proofs of the Theorems

Proof of Theorem 3.1. Map $(\Omega, (\Pi^i)_{i\in I}, f : A \to 2^\Omega)$ to a subspace of W_α by H_α, where α is large enough to make it an epimorphism by theorem 4.1. Let $\omega \in \Omega$, $w_\alpha = H_\alpha(\omega)$ and $\Gamma \subseteq \mathcal{L}_{c(\alpha)}$ the set of formulas that hold in w_α. Since an epimorphism preserves truth of formulas, Γ is also the set of formulas of $\mathcal{L}_{c(\alpha)}$ that obtain in ω. Now, let $\widehat{W}_\alpha$ be the minimal subspace of W_α that contains w_α (It is the intersection of all subspaces that contain w_α). H_α on $(\widehat{W}_\alpha, (\Pi^i_\alpha)_{i\in I}, f_\alpha : A \to 2^{\widehat{W}_\alpha})$ is the identity, as can be easily verified. Therefore, by theorem 4.1, $\widehat{W}_\alpha$ is non-redundant and minimal. Had there been another model with these properties, we could have embedded it similarly as a subspace of some $W_{\alpha'}$. As every state in every subspace has a unique extension to higher-level W_β-s, the extensions of the two subspaces to W_β where $\beta \geq \max(\alpha, \alpha')$ would be the same. Hence $\widehat{W}_\alpha$ is unique (up to isomorphism), as required. ■

To prove theorem 2.1 we first need the following propositions:

Proposition 5.1. $\Gamma \models \varphi$ in $\mathcal{L}_c$ if and only if φ holds in all the states of W_α where Γ holds, where $\alpha = \text{rank}(\Gamma \cup \{\varphi\})$.

Proposition 5.2. For every ordinal α there is a cardinal $c(\alpha)$, such that for every $w_\alpha \in W_\alpha$ there is a formula $\varphi(w_\alpha) \in \mathcal{L}_{c(\alpha)}$ with the following properties:

(1) $\varphi(w_\alpha)$ holds in w_α.

(2) For every $\psi \in \mathcal{L}_{c(\alpha)}$ with $\text{rank}(\psi) \leq \alpha$, either $\varphi(w_\alpha) \vdash \psi$ or $\varphi(w_\alpha) \vdash \neg\psi$.

(3) $\vdash \bigvee_{w_\alpha \in W_\alpha} \varphi(w_\alpha)$.

Once we have proved these propositions, the proof of theorem 2.1 is straightforward:

Proof of Theorem 2.1. Suppose $\Gamma \models \varphi$ in $\mathcal{L}_c$, where $c \geq c(\alpha)$ from proposition 5.2 and $\alpha = \text{rank}(\Gamma \cup \{\varphi\})$. By proposition 5.1 this happens if and only if φ holds in all the states of W_α where Γ holds. By (2) from proposition 5.2 we must have $\vdash \varphi(w_\alpha) \to \varphi$ for every w_α where Γ holds, so also

$$(*) \qquad \vdash \left(\bigvee\{\varphi(w_\alpha) : \Gamma \text{ holds in } w_\alpha\}\right) \to \varphi.$$

On the other hand, for every w_α where Γ does not hold we must have $\vdash \varphi(w_\alpha) \to \neg\psi$ for some $\psi \in \Gamma$, and hence

$$\Gamma \vdash \bigwedge\{\neg\varphi(w_\alpha) : \Gamma \text{ does not hold in } w_\alpha\}.$$

By (3) of 5.2 we therefore have

$$\Gamma \vdash \bigvee\{(w_\alpha) : \Gamma \text{ holds in } w_\alpha\},$$

so by $(*)$ $\Gamma \vdash \varphi$. ■

It remains to prove the two propositions.

Proof of proposition 5.1. It is easy to verify, by transfinite induction on the ranks of formulas, that if $\alpha = \text{rank}(\psi)$ and ψ holds in $w_\alpha \in W_\alpha$, then ψ also holds in any extension of w_α to a state w_β for $\beta > \alpha$. (This follows from property (III) in the definition of W_α.) Therefore, for $\alpha = \text{rank}(\Gamma \cup \{\varphi\})$, if φ holds wherever Γ holds in W_α, then this obtains also in any W_β, $\beta > \alpha$. By theorem 4.1 any model of $\mathcal{L}_c$ is epimorphic to a subspace of W_γ for some $\gamma \geq \alpha$, and an epimorphism preserves the truth of formulas. Therefore, if φ holds wherever Γ holds in W_α then $\Gamma \models \varphi$. The implication in the other direction is simply due to the fact that W_α is one particular model. ■

Proof of proposition 5.2. Define, by transfinite induction, the following formulas $\varphi(t^i_\gamma)$ and $\varphi(w_\alpha)$ for all $w_\alpha = (w_o, (t^i_\gamma)_{\substack{\gamma<\alpha \\ i\in I}}) \in W_\alpha$:

$$\varphi(w_0) = \bigwedge_{\varphi \in w_0} \varphi \wedge \bigwedge_{\varphi \notin w_0} \neg\varphi$$

(Recall that $W_0 = 2^A$),

$$\varphi(t^i_\gamma) = \bigwedge_{t^i_\gamma \subseteq V \subseteq W_\gamma} k^i\Big(\bigvee_{w_\gamma \in V} \varphi(w_\gamma)\Big) \wedge \bigwedge_{t^i_\gamma \not\subseteq V \subseteq W_\gamma} \neg k^i\Big(\bigvee_{w_\gamma \in V} \varphi(w_\gamma)\Big)$$

$$\varphi(w_\alpha) = \varphi(w_0) \wedge \bigwedge_{\substack{\gamma<\alpha \\ i\in I}} \varphi(t^i_\gamma)$$

Choose $c(0) > 2^{|A|}$ and $c(\alpha) > \sup\limits_{\beta<\alpha}(2^{2^{|W_\beta|}}) \times |I|$ for $\alpha > 0$. Then clearly $\varphi(w_\alpha) \in \mathcal{L}_{c(\alpha)}$. Now, articles (1) and (2) of the proposition can be easily verified by induction on the structure of formulas up to rank α, relying on the coherence property (III) of the t^i_γ in the definition of W_α. It remains to prove (3).

For $\alpha = 0$ we have $\vdash \varphi \vee \neg\varphi$ for every atomic $\varphi \in A$, so by axiom (A6)

$$\vdash \bigvee_{w_0 \in 2^A} \big(\bigwedge_{\varphi \in w_0} \varphi \wedge \bigwedge_{\varphi \notin w_0} \neg\varphi \big)$$

which is exactly $\bigvee_{w_0 \in W_0} \varphi(w_0)$.

Suppose we have proved (3) for all $\beta < \alpha$. Fix $i \in I$ and $\beta < \alpha$ for the discussion below. Then we have

$$(i) \qquad \vdash \varphi(w_\beta) \rightarrow \neg k^i \neg \varphi(w_\beta)$$

(by (K1)), and

$$(ii) \qquad \vdash \varphi(w_\beta) \rightarrow \varphi(t^i_\gamma) \qquad \forall \gamma < \beta$$

(by the definition of $\varphi(w_\beta)$). Since $\varphi(t^i_\gamma)$ is a conjunction of formulas that start with k^i or $\neg k^i$, by (K3), (K4) and (K2) we have $\vdash \varphi(t^i_\gamma) \rightarrow k^i \varphi(t^i_\gamma)$, so also

$$(iii) \qquad \vdash \varphi(w_\beta) \rightarrow k^i \varphi(t^i_\gamma) \qquad \forall \gamma < \beta.$$

Now, for each $V \subseteq W_\beta$ let

$$\varphi(V) = k^i \big(\bigvee_{w_\beta \in V} \varphi(w_\beta) \big).$$

Then we clearly have $\vdash \varphi(V) \vee \neg\varphi(V)$, and by conjunction (over all $V \subseteq W_\beta$) and (A6) we get

$$(*) \qquad \vdash \bigvee_{X \subseteq 2^{W_\beta}} \big(\bigwedge_{V \in X} \varphi(V) \wedge \bigwedge_{V \notin X} \neg\varphi(V) \big)$$

(Notice that even this huge disjunction is still a formula in $\mathcal{L}_{c(\alpha)}$.)

However, when X is not closed under supersets, the disjunct for X would be contradictory, because there would be $V \in X$, $V \subseteq V' \notin X$, and the disjunct for X would imply

$$k^i \big(\bigvee_{w_\beta \in V} \varphi(w_\beta) \big) \wedge \neg k^i \big(\bigvee_{w_\beta \in V'} \varphi(w_\beta) \big),$$

contradicting the monotonicity axiom (K2).

Furthermore, when $\bigcap_{V \in X} V \notin X$, the disjunct for X in $(*)$ would also be contradictory, because

$$\vdash \bigwedge_{V \in X} \big(\bigvee_{w_\beta \in V} \varphi(w_\beta) \big) \rightarrow \bigvee_{w_\beta \in \bigcap_{V \in X} V} \varphi(w_\beta),$$

so by (K2)

$$\bigwedge_{V\in X} \varphi(V) \wedge \neg\varphi(\bigcap_{V\in X} V)$$

is contradictory, while it is implied from the disjunct for X.

This means that in $(*)$ X may be limited to be closed under supersets and to have a least element. If for $V_0 \subseteq W_\beta$ we denote

$$\varphi(V_0) = \bigwedge_{V_0\subseteq V} \varphi(V) \wedge \bigwedge_{V_0\not\subseteq V} \neg\varphi(V)$$

then we conclude

$$(**) \qquad \vdash \bigvee_{V_0\subseteq W_\beta} \varphi(V_0).$$

Fix $(w_0,\ (t^i_\gamma)_{\substack{\gamma<\beta\\ i\in I}}) \in W_\beta$. If V_0 is the t^i_β-coordinate in some extension of w_β to a state in $W_{\beta+1}$, then (by definition) $\varphi(V_0) = \varphi(t^i_\beta)$. Otherwise, one of the conditions in the definition of W_α are violated by V_0, so either

$$\text{(I)} \qquad \vdash \varphi(V_0) \to k^i\neg\varphi(w_\beta)$$

or

$$\text{(II)} \qquad \vdash \varphi(V_0) \to \neg k^i\varphi(t^i_\gamma) \qquad \text{for some } \gamma < \beta$$

or

$$\text{(III)} \qquad \vdash \varphi(V_0) \to \neg\varphi(t^i_\gamma) \qquad \text{for some } \gamma < \beta.$$

From (i), (ii), (iii) and $(**)$ we now conclude

$\vdash \varphi(w_\beta) \to \bigvee\{\varphi(V_0) : V_0$ is the t^i_β-coordinate in an extension of w_β to a state in $W_{\beta+1}\}$

By conjunction over all $i \in I$ and axiom (A6) we get

$$\vdash \varphi(w_\beta) \to \bigvee\{\bigwedge_{i\in I} \varphi(t^i_\beta) :\ (w_\beta,\ (t^i_\beta)_{i\in I}) \in W_{\beta+1}\}.$$

From the induction hypothesis we already have

$$\vdash \bigvee_{w_\beta\in W_\beta} \varphi(w_\beta),$$

and together we get

$$\vdash \bigvee\{\varphi(w_\beta) \wedge \bigwedge_{i\in I} \varphi(t^i_\beta) :\ (w_\beta,\ (t^i_\beta)_{i\in I}) \in W_{\beta+1}\}$$

or, in other words,

$$\vdash \bigvee\{\varphi(w_0) \wedge \bigwedge_{\substack{\gamma<\beta+1\\ i\in I}} \varphi(t^i_\gamma) :\ (w_0,\ (t^i_\gamma)_{\substack{\gamma<\beta+1\\ i\in I}}) \in W_{\beta+1}\}$$

Now, by Conjunction over all $\beta < \alpha$ and by activating (A6) again we finally get

$$\vdash \bigvee\{\varphi(w_0) \wedge \bigwedge_{\substack{\gamma<\alpha\\ i\in I}} \varphi(t^i_\gamma) :\ (w_0,\ (t^i_\gamma)_{\substack{\gamma<\alpha\\ i\in I}}) \in W_\alpha\}$$

which is the desired conclusion. ∎

References

Aumann, R.(1976), Agreeing to Disagree, Annals of Statistics 4, pp. 1236-1239.

Aumann, R.(1989), Notes on Interactive Epistemology, Cowles Foundation for Research in Economics working paper.

Chellas, B.F. (1980), Modal Logic, an Introduction, Cambridge University Press.

Fagin, R.(1993), A Quantitative Analysis of Modal Logic, forthcoming in JSL.

Fagin, R., Geanakoplos, J., Halpern, J. Y. and Vardi, M. Y.(1992), The Expressive Power of the Hierarchical Approach to Modeling Knowledge and Common Knowledge, Proc. of TARK IV, pp. 229-244.

Fagin, R., Halpern, Y.J. and Vardi, M.Y.(1991), A Model-Theoretic Analysis of Knowledge, Jour. of the ACM 91(2), pp. 382-428.

Heifetz, A.(1992), Hierarchic Constructions of Incomplete Information Models, preprint, School of Mathematical Sciences, Tel Aviv University.

Heifetz, A. and Samet, D.(1993), Universal Partition Structures, IIBR working paper 26/93, Tel Aviv University.

Hughes G.E. and Cresswell, M.J. (1984), A companion to Modal Logic, Methuen, London.

Karp, C. R.(1964), Languages with Expressions of Infinite Length, North Holland, Amsterdam.