

# The Expressive Power of The Hierarchical Approach to Modeling Knowledge and Common Knowledge

Ronald Fagin  
 IBM Almaden Research Center  
 San Jose, CA 95120  
 email: fagin@almaden.ibm.com

John Geanakoplos  
 Cowles Foundation, Yale University  
 New Haven, CT 06520

Joseph Y. Halpern  
 IBM Almaden Research Center  
 San Jose, CA 95120  
 email: halpern@almaden.ibm.com

Moshe Y. Vardi  
 IBM Almaden Research Center  
 San Jose, CA 95120  
 email: vardi@almaden.ibm.com

## Abstract

One approach to representing knowledge or belief of agents, which has been explored independently by economists (Böge and Eisele; Mertens and Zamir; Brandenburger and Dekel; Tan and Werlang) and by computer scientists (Fagin, Halpern, and Vardi) involves an infinite hierarchy of beliefs. Such a hierarchy consists of an agent's beliefs about the state of the world, his beliefs about other agents' beliefs about the worlds, his beliefs about other agents' beliefs about other agents' beliefs about the worlds, etc. Economists and computer scientists differ, however, in the way they model beliefs. Economists prefer a probability-based framework, where belief is modeled as a probability distribution on the uncertainty space. In contrast, computer scientists prefer an information-based framework, where belief is modeled as a subset of the underlying space. The idea is that whatever is in the subset is believed to be possible, and whatever is not in the subset is believed to be impossible.

We consider the question of when such an infinite hierarchy completely describes the uncertainty of the agents. We provide various necessary and sufficient conditions for this property. It turns out that the probability-based approach can be viewed as satisfying one of these conditions, which explains why the infinite hierarchy always completely describes the uncertainty of the agents in the probability-based approach. An interesting consequence of our conditions is that adequacy of an infinite hierarchy may depend on the "richness" of the states in the underlying state space. We also consider the question of whether an infinite hierarchy completely describes the uncertainty of the agents with respect to "interesting" sets of events and show that the answers depends on the definition of "interesting".

## 1 Introduction

Reasoning about knowledge of agents and their knowledge of each other's knowledge has now been recognized as a fundamental concern in game theory, computer science, artificial intelligence, and philosophy (see [Hal91] for a recent overview). The importance of finding good formal models that can represent the knowledge of the agents has also been long recognized.

The original approach to representing knowledge and common knowledge in the game-theory literature is due to Aumann [Aum76]. Consider a situation with  $n$  agents. To model this, Aumann considers structures of the form  $A = (S, \mathcal{K}_1, \dots, \mathcal{K}_n)$ , where  $S$  is a set of states of the world, and each  $\mathcal{K}_i$  is a partition of  $S$ . We henceforth call these *Aumann structures*.<sup>1</sup> An agent “knows” about events, which are identified with subsets of  $S$ . Agent  $i$ ’s knowledge is modeled by  $\mathcal{K}_i$ , his *information partition*. Given a state  $s \in S$ , we use  $\mathcal{K}_i(s)$  to denote the set of states in the same element of the partition as  $s$ ; these are the states that agent  $i$  considers to be possible in state  $s$ . Agent  $i$  is said to *know* an event  $E$  at the state  $s$  if  $\mathcal{K}_i(s)$  is a subset of  $E$ . The intuition behind this is that in state  $s$ , agent  $i$  cannot distinguish between any of the worlds in  $\mathcal{K}_i(s)$ . Thus, agent  $i$  knows  $E$  in state  $s$  if  $E$  holds at all the states that  $i$  cannot distinguish from  $s$ . Using this intuition, we can define an operator  $K_i$  from events to events. Given an event  $E$ , the event  $K_i(E)$  (intuitively, the event “agent  $i$  knows  $E$ ”) can be identified with the set of states where agent  $i$  knows  $E$  according to our definition. We can also define the event  $O(E)$  (“everyone knows  $E$ ”) as the intersection of the events  $K_i(E)$ , over all agents  $i = 1, \dots, n$ . Finally, we can define the event  $C(E)$  (“ $E$  is common knowledge”) as the intersection of the events  $O(E)$ ,  $O(O(E))$ , and so on.

There is, unfortunately, a philosophical difficulty with this approach (cf. [Gil88, TW88, Aum89]). The problem is that it is not *a priori* clear what the relation is between a state in an Aumann structure—which is, after all, just an element of a set—and the rather complicated reality that this state is trying to model. If we think of a state as a complete description of the world, then it must capture all of the agents’ knowledge. Since the agents’ knowledge is defined in terms of the partitions, the state must include a description of the partitions. This seems to lead to circularity, since the partitions are defined over the states, but the states contains a description of the partitions.

Partly in response to these concerns, an alternative approach to modeling knowledge was investigated in a number of papers [BE79, MZ85, TW88, BD92]. This approach, which involves an infinite hierarchy of beliefs, takes its cue from the work of Harsanyi [Har68]. We start with a set  $S$  of *states of nature*, which we take to be descriptions of certain facts about the world (e.g., in game theory, these may typically include values of parameters of the game, such as payoffs). Each agent has beliefs about the state of nature, modeled by a probability distribution over  $S$ . These beliefs are clearly highly relevant to the agent’s choice of strategy. But agents also have beliefs about other agents’ beliefs, and beliefs about other agents’ beliefs about their beliefs, and so on. Pursuing this line, one is naturally led to associating with each agent a hierarchy of beliefs. We can build up this hierarchy level by level; for each natural number  $m$ , the  $(m + 1)$ <sup>st</sup>-order beliefs of agent  $i$  are modeled by a probability distribution on the possible states of nature and the other agents’  $m$ <sup>th</sup>-order beliefs (together with some consistency conditions described in [MZ85, BD92]). An agent’s *type* is his infinite hierarchy of beliefs. We define a *belief structure* to consist of a state of nature and a description of each agent’s type. Given a set  $S$  of states of the world, we take  $\mathcal{B}(S)$  to be the set of belief structures where  $S$  is the set of states of nature.

In belief structures, knowledge is identified with “believes with probability 1”. That is, roughly speaking, agent  $i$  is said to know  $E \subseteq S$  in a given belief structure  $\mathbf{b}$  if, according to agent  $i$ ’s type in  $\mathbf{b}$ , event  $E$  is assigned probability 1 at level 1 of agent  $i$ ’s hierarchy. Similarly, agent  $i$  knows that agent  $j$  knows  $E$  if the event “agent  $j$  knows  $E$ ” is assigned probability 1 at level 2 of agent

<sup>1</sup>The reader with a background in modal logic will recognize that an Aumann structure is nothing more than a *Kripke frame for  $S5$*  [Kri59, HC78, HM92]. In Aumann’s original paper, he also assumed that there was a probability distribution on  $S$ . Since the probability function plays no role in our discussion of knowledge in Aumann structures, we have decided to drop it here. This is consistent with Aumann’s own discussion of knowledge in later papers (see, for example, [Aum89]) and with the presentation of Aumann’s framework in, for example, [Wer89].

$i$ 's type hierarchy. Finally, we say that  $E$  is common knowledge if all agents know  $E$ , all agents know they know  $E$ , and so on.

We would like to think of a belief structure as describing a state of the world. It is not clear, however, that a belief structure is an adequate description of the state of the world. Even if we accept the doctrine that a state of the world can be adequately described by describing the actual state of nature and each agent's uncertainty about the state of nature and other agents' uncertainty (at all levels), it is not clear that the infinite hierarchy just described completely exhausts an agent's uncertainty. After all, an agent may have uncertainty as to the type of other agents. Harsanyi essentially assumed that there is an exogenously given probability distribution that describes each agent's probability distribution on the state of nature and the other agents' types. The key result proved in [BE79, MZ85] is that the hierarchy described above does exhaust an agent's beliefs: an agent's type determines a unique probability distribution on the states of nature and the other agents' types.

This result also suggests that we can view the belief structures in  $\mathcal{B}(S)$  as the states in an Aumann structure, since each one completely describes a state of the world. If we take that view, then we might hope that the definitions of knowledge and common knowledge in Aumann structures and belief structures coincide. Unfortunately, this is not quite the case. Nevertheless, Brandenburger and Dekel [BD92] show that these notions do coincide if we interpret knowledge in Aumann structures probabilistically. Thus, we view  $\mathcal{B}(S)$  as an Aumann structure, with the information partitions being determined by the type (so that two belief structures  $\mathbf{b}$  and  $\mathbf{b}'$  are in the same equivalence class of  $\mathcal{K}_i$  iff agent  $i$  has the same type in  $\mathbf{b}$  and  $\mathbf{b}'$ ). In addition, we endow  $\mathcal{B}(S)$  with probability measures  $\mu_i$  (one probability for each agent  $i$ ) based on information in the individual belief structures (for more details on the construction, see [BD92]). We then take the event "agent  $i$  knows  $E$ " to hold in state  $s$  if  $\mu_i(E|\mathcal{K}_i(s)) = 1$ ; similar modifications are necessary for common knowledge. (We identify the event  $E \subseteq S$  with the subset of  $\mathcal{B}(S)$  consisting of all belief structures for which the state of nature is  $E$ .) Brandenburger and Dekel then show that an event  $E \subseteq S$  is common knowledge in a state  $\mathbf{b}$  in the (probabilistically endowed) Aumann structure  $\mathcal{B}(S)$  iff  $E$  is common knowledge in the belief structure  $\mathbf{b}$ . A complementary result is proved in [TW88], where it is shown that given an Aumann structure  $A = (S, \mathcal{K}_1, \dots, \mathcal{K}_n)$  and  $s_0 \in S$ , there is a belief structure  $\mathbf{b} \in \mathcal{B}(S)$  such that an event  $E \subseteq S$  is common knowledge at  $s_0$  iff  $E$  is common knowledge in  $\mathbf{b}$ .

This may seem to pretty much complete the picture: the hierarchical approach provides the answer to the problem of circularity in Aumann structures, since the above results seem to indicate that belief structures are adequate in modeling the states in Aumann structures. Unfortunately, the situation is somewhat more complicated than these results suggest. The fundamental problem with these results is that they are trying to relate two incomparable concepts of knowledge: the information-theoretic concept in Aumann structures and the probability-theoretic concept in belief structures (which is why Brandenburger and Dekel had to recast Aumann's framework in a probabilistic setting). We would argue that the probabilistic framework masks some of the subtleties in the issue of the adequacy of the hierarchical approach. Since the circularity issue in Aumann structures arises in a non-probabilistic setting, we believe that the adequacy of the hierarchical approach is best examined in a non-probabilistic setting.

A non-probabilistic setting for the hierarchical approach is described in [FHV91]. We again start with states of nature, and build a hierarchy, level by level. In this case, the  $(m + 1)$ <sup>st</sup>-order knowledge of agent  $i$  is modeled by a set of possibilities, each of which is a description of a state of nature and each agent's  $m$ <sup>th</sup>-order knowledge (satisfying certain consistency conditions).

Intuitively, whatever is in the subset is believed to be possible, and whatever is not in the subset is believed to be impossible. Note that there is no probability distribution, just a set of possibilities. A *knowledge structure* consists of a state of nature and each agent's hierarchy of possibilities. We take  $\mathcal{F}(S)$  to be the set of knowledge structures where  $S$  is the set of states of nature.

Knowledge and common knowledge are defined in knowledge structures in an information-theoretic fashion, as in Aumann structures. That is, agent  $i$  is said to know  $E \subseteq S$  in a given knowledge structure if the set of states that  $i$  considers possible at level 1 is a subset of  $E$ ; agent  $j$  knows that agent  $i$  knows  $E$  if the set of sequences of length 2 that  $j$  considers possible at level 2 is a subset of the set of sequences of length 2 where  $i$  knows  $E$ . Common knowledge is again defined in the standard way in terms of knowledge.

In [FHV91], results connecting knowledge structures and Aumann structures analogous to those of [BD92] and [TW88] are proved. Namely, it is shown that we can view  $\mathcal{F}(S)$  as an Aumann structure, where the partitions are determined by the agents' types, and an event  $E \subseteq S$  is common knowledge in a knowledge structure  $f \in \mathcal{F}(S)$  according to Aumann's definition iff  $E$  is common knowledge at  $f$  according to the knowledge structures definition. Moreover, it is shown that given an Aumann structure  $A = (S, \mathcal{K}_1, \dots, \mathcal{K}_n)$  and a state  $s_0 \in S$ , there is a knowledge structure  $f \in \mathcal{F}(S)$  such that an event  $E \subseteq S$  is common knowledge at  $s_0$  iff  $E$  is common knowledge in  $f$ .

This seems to confirm the results of [TW88, BD92] and suggest that the hierarchical approach does address the circularity problem. Unfortunately, it is also shown in [FHV91] that knowledge structures are in general *not* an adequate description of the world, since they do not completely describe an agent's uncertainty. In particular, an agent's type does not determine what other types the agent considers possible. The problem is that the hierarchy in knowledge structures (as well as in belief structures) contain only  $\omega$  levels, when in general we need to consider transfinite hierarchies.

There seems to be an inconsistency here. How can it be the case that knowledge structures are not an adequate description of an agent's knowledge while belief structures are, and how we do reconcile the inadequacy of knowledge structures with the results relating knowledge structures to Aumann structures? Our goal in this paper is to address these questions by examining the adequacy of hierarchical structures and trying to make precise how expressive they are. We do this in the non-probabilistic framework of knowledge structures, but we go beyond [FHV91] in our focus on the adequacy issue, and in our logic-free setting.

We start by considering more carefully the question of when a knowledge structure does completely characterize the agents' knowledge. We provide a necessary and sufficient condition for this property. A surprising consequence of our condition is that a knowledge structure completely characterizes the agents' knowledge if and only if it characterizes the first  $\omega + \omega$  levels of knowledge. Another consequence of our condition is that the adequacy of knowledge structures may depend on the "richness" of the states in the underlying state space  $S$ . If the states of nature are modeled in enough detail, then knowledge structures do characterize the agents' knowledge.

Our analysis also shows the importance of a certain *limit-closure* property, which says that what happens at finite levels determines what happens at the limit. This property holds for belief structures, because of the assumption that probability is countably additive. The fact that this property holds for belief structures is precisely why the result of [BE79, MZ85] holds. If we allowed finitely additive probabilities, then belief structures would not, in general, completely characterize the agents' beliefs.

Since knowledge structures do not, in general, characterize agents' knowledge, it is inappropri-

ate here to study their relationship to Aumann structures by simply defining partitions on  $\mathcal{F}(S)$ . We examine this relationship in a more general setting, by asking whether knowledge structures characterize agent's knowledge with respect to "interesting" sets of events. The answer, of course, depends on what is considered to be an "interesting" set of events. It turns out, for example, that if we consider only events that can be defined from "natural events" by knowledge and common knowledge operators, then knowledge structures are adequate. If, on the other hand, we are interested in common knowledge among coalitions of agents (rather than just common knowledge among *all* the agents), then knowledge structures are not adequate. In this case, a transfinite hierarchy is necessary, but  $\omega^2$  levels suffice.

This discussion gives the impression that the only issue underlying the adequacy of the hierarchical approach is that of the "length" of the hierarchy. But it is easy to see that knowledge structures are also deficient in a way that no transfinite hierarchy can remedy. Aumann structures contain information about all *conceivable* states, even states that are commonly known not to hold. Thus, Aumann structures enable *counterfactual* reasoning, such as "If Ron Fagin were the President, then he would not have stopped the war against Iraq so soon." A counterfactual statement can be viewed as a statement about a world commonly known not to be possible. (It is presumably common knowledge that Ron Fagin is not the President.) Knowledge structures, on the other hand, do not enable such reasoning, since situations commonly known to be impossible never appear as prefixes in knowledge structures.

It turns out that this deficiency is not inherent in the hierarchical approach, but rather it is the result of the manner in which this approach was used in knowledge structures. Knowledge structures were designed to model knowledge; no more, no less. As we show, the hierarchical approach can also be used to define structures that do capture information about conceivable states. These results suggest that hierarchical structures can always serve as adequate models of the world. In general, however, we may need to capture more than just knowledge and we may need to continue the hierarchy into the transfinite ordinals, in order to completely capture the agents' uncertainty. What we choose to capture and how far into the ordinals we need to go depends on the events that we are interested in capturing. Thus, the question of whether knowledge or belief structures as defined are adequate models depends on both what features of the world we are trying to model, and on the events we are interested in describing.

## 2 Knowledge structures and belief structures: a review

In this section we review the definitions of knowledge structures and belief structures. We begin with knowledge structures. The following material is largely taken from [FHV91], slightly modified to be consistent with the rest of our presentation here.

We start with a set  $S$  of states of nature, a fixed finite set  $\{1, \dots, n\}$  of agents. We now define  $k$ -ary worlds, by induction on  $k$ . A 0th-order knowledge assignment,  $f_0$ , is a member of  $S$ , that is, a state of nature (which, intuitively, corresponds to the "real world"). We call  $\langle f_0 \rangle$  a 1-ary world (since its length is 1). Assume inductively that  $k$ -ary worlds (or  $k$ -worlds, for short) have been defined. Let  $W_k$  be the set of all  $k$ -worlds. A  $k$ th-order knowledge assignment (for  $k \geq 1$ ) is a function that associates with each agent  $i$  a set  $f_k(i) \subseteq W_k$  of "possible  $k$ -worlds"; we think of the worlds in  $f_k(i)$  as "possible" for agent  $i$  and the worlds in  $W_k - f_k(i)$  as "impossible" for agent  $i$ . A  $(k + 1)$ -sequence of knowledge assignments is a sequence  $\langle f_0, \dots, f_k \rangle$ , where  $f_i$  is an  $i$ th-order knowledge assignment. A  $(k + 1)$ -world is a  $(k + 1)$ -sequence of knowledge assignments

that satisfies certain conditions described below. These conditions enforce some intuitive properties of knowledge. An infinite sequence  $\langle f_0, f_1, f_2, \dots \rangle$  is called a *knowledge structure* if for each  $k$ , each *prefix*  $\langle f_0, \dots, f_{k-1} \rangle$  is a  $k$ -world. Thus, a  $k$ -world describes knowledge of depth  $k - 1$ , and a knowledge structure describes knowledge of arbitrary depth. We use  $\mathcal{F}(S)$  to denote the set of knowledge structures over  $S$ .

A  $(k+1)$ -world  $\langle f_0, \dots, f_k \rangle$  must satisfy the following restrictions for each agent  $i$ :

- (K1) **Correctness:**  $\langle f_0, \dots, f_{k-1} \rangle \in f_k(i)$ , if  $k \geq 1$ . Intuitively, this condition says that knowledge is always correct (unlike belief, which can be incorrect).
- (K2) **Introspection:** If  $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$ , and  $k > 1$ , then  $g_{k-1}(i) = f_{k-1}(i)$ . This condition implies that our agents are introspective about their own knowledge; at each level  $k$ , they know exactly what they know and what they don't know at level  $k - 1$ .
- (K3) **Extension:**  $\langle g_0, \dots, g_{k-2} \rangle \in f_{k-1}(i)$  iff there is a  $(k - 1)$ <sup>st</sup>-order knowledge assignment  $g_{k-1}$  such that  $\langle g_0, \dots, g_{k-2}, g_{k-1} \rangle \in f_k(i)$ , if  $k > 1$ . Intuitively, this condition says that the different levels of knowledge describing a knowledge world are consistent with each other.

Let  $\mathbf{f}$  be the knowledge structure  $\langle f_0, f_1, \dots \rangle$ . Define  $i$ 's *view of  $\mathbf{f}$* , or  $i$ 's *type in  $\mathbf{f}$* , denoted  $\pi_i(\mathbf{f})$ , to be the sequence  $\langle f_1(i), f_2(i), \dots \rangle$ . We write  $\mathbf{f} \sim_i \mathbf{f}'$  if  $\pi_i(\mathbf{f}) = \pi_i(\mathbf{f}')$ , that is, if  $i$  has the same type in  $\mathbf{f}$  and  $\mathbf{f}'$ .

We can now ask whether knowledge structures as defined above are adequate to fully capture all of an agent's knowledge. As we shall discuss shortly, it is shown in [FHV91] that they are *not* adequate. That is, an agent's type does not necessarily determine the set of knowledge structures that he considers possible. At this point, it is not even clear what we mean by "the knowledge structures that agent  $i$  considers possible". In a given knowledge structure  $\mathbf{f} = \langle f_0, f_1, \dots \rangle$ , it should be clear, for each finite  $k$ , what we mean by "the  $k$ -worlds that agent  $i$  considers possible": these are simply the worlds in  $f_k(i)$ . In the case of knowledge structures, rather than  $k$ -worlds, we would similarly like to say that agent  $i$  considers the knowledge structure  $\mathbf{g}$  possible in  $\mathbf{f}$  precisely if  $\mathbf{g} \in f_\omega(i)$ . In fact, this is exactly what we shall do shortly, when we extend the hierarchy into the transfinite ordinals. There is, however, no  $f_\omega$  in a knowledge structure. So how do we make sense of "the knowledge structures that agent  $i$  considers possible in  $\mathbf{f}$ "? One way is to say that agent  $i$  considers the knowledge structure  $\mathbf{g}$  possible in  $\mathbf{f}$  precisely if  $\mathbf{f} \sim_i \mathbf{g}$ . This says that agent  $i$  considers the knowledge structure  $\mathbf{g}$  possible precisely if agent  $i$  has the same type in  $\mathbf{g}$  as in  $\mathbf{f}$ . Another way to say that agent  $i$  considers the knowledge structure  $\mathbf{g}$  possible in  $\mathbf{f}$  is if agent  $i$  considers every finite prefix of  $\mathbf{g}$  possible. This second notion says that if  $\mathbf{g} = \langle g_0, g_1, \dots \rangle$ , then agent  $i$  considers the knowledge structure  $\mathbf{g}$  possible in  $\mathbf{f}$  precisely if  $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$  for every  $k \geq 1$ . The next theorem says that these two notions are equivalent.

**Theorem 2.1:** [FHV91] *Let  $\mathbf{f} = \langle f_0, f_1, \dots \rangle$  and  $\mathbf{g} = \langle g_0, g_1, \dots \rangle$  be knowledge structures. Then  $\mathbf{f} \sim_i \mathbf{g}$  iff  $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$  for every  $k \geq 1$ .*

We can think of the first notion in Theorem 2.1 (that is,  $\mathbf{f} \sim_i \mathbf{g}$ ) as an *external* notion of possibility. There can be uncountably many knowledge structures  $\mathbf{g}$  such that  $\mathbf{f} \sim_i \mathbf{g}$ . The other notion is *internal*: we consider every  $k$ -world that  $i$  considers possible, by "looking inside" the knowledge structure (at level  $k$ ). Thus, this second notion is finitistic. Theorem 2.1 tells us that the external and internal notions coincide.

As we hinted above, to understand whether knowledge structures fully capture the agents' knowledge, we need to continue the construction of the hierarchy into the transfinite ordinals. We can view knowledge structures as we have defined them as  $\omega$ -worlds. To construct  $(\omega + 1)$ -worlds, we need a function that tells us, for each agent  $i$ , which  $\omega$ -worlds  $i$  considers possible. We can then continue on inductively to define  $\lambda$ -worlds for every ordinal  $\lambda$ . Suppose we have defined  $\lambda$ -worlds for some ordinal  $\lambda$ . Let  $W_\lambda$  be the set of all  $\lambda$ -worlds. A  $\lambda^{\text{th}}$ -order knowledge assignment is a function that associates with each agent  $i$  a set  $f_\lambda(i) \subseteq W_\lambda$ . A  $\lambda$ -sequence of knowledge assignments is a sequence  $\langle f_0, f_1, \dots \rangle$  of length  $\lambda$ , where  $f_i$  is an  $i^{\text{th}}$ -order knowledge assignment. A  $\lambda$ -world  $f$  is a  $\lambda$ -sequence of knowledge assignments satisfying conditions that are straightforward extensions of conditions K1–K3 above. For example, an  $(\omega + 1)$ -world is of the form  $\langle f_0, f_1, \dots, f_\omega \rangle$ , where  $f_\omega(i)$  is a set of  $\omega$ -worlds satisfying the appropriate conditions. In the sequel, when we speak of knowledge structures, we will mean  $\omega$ -worlds.

Consider now an  $(\omega + 1)$ -world  $f = \langle f_0, f_1, \dots, f_\omega \rangle$ . There are now two ways we can define “the knowledge structures that agent  $i$  considers possible in  $f$ ”. One way is to say that agent  $i$  considers the knowledge structure  $g$  possible in  $f$  precisely if  $g \in f_\omega(i)$ . Note, however, that the first  $\omega$  levels constitute an  $\omega$ -world  $f' = \langle f_0, f_1, \dots \rangle$ . Thus, another way is to say that agent  $i$  considers the knowledge structure  $g$  possible in  $f$  precisely if agent  $i$  considers the knowledge structure  $g$  possible in  $f'$ , i.e., precisely if  $f' \sim_i g$ . It is shown in [FHV91] that these two ways are *not* equivalent; while we have that  $f_\omega(i) \subseteq \{g \mid f' \sim_i g\}$ , equality need not hold. Thus, knowledge structures do not fully describe the agents' knowledge.

Belief structures are defined along similar lines. We briefly sketch the definition here, and refer the reader to [MZ85, TW88, BD92] for more details. We start with  $S$ , which we assume is endowed with a topology that makes it a compact metric space.<sup>2</sup> Given a compact metric space  $X$ , let  $\Delta(X)$  denote the set of Borel probability measures on  $X$ . If we endow  $\Delta(X)$  with the topology of weak convergence of measures, then  $\Delta(X)$  is also a compact metric space. Define a sequence of spaces  $X_k, k = 0, 1, 2, \dots$  inductively by taking  $X_0 = S$  and

$$X_{k+1} = X_k \times \Delta(X_k)^n (= X_0 \times \Delta(X_0)^n \times \Delta(X_1)^n \times \dots \times \Delta(X_k)^n).$$

A belief structure  $b$  is a sequence  $\langle b_0, b_1, \dots \rangle$  such that  $b_0 \in S$ , and  $b_k \in \Delta(X_{k-1})^n$  for each  $k > 0$ . This means that for  $k > 0$  we can view  $b_k$  as a function such that for each agent  $i$ , we have  $b_k(i) \in \Delta(X_{k-1})$ . We have consistency conditions B1 and B2 on belief structures that correspond to K2 and K3:

- (B1) For all  $k > 1$ , the probability measure  $b_k(i)$  assigns probability 1 to the subspace of  $X_{k-1}$  consisting of sequences  $\langle c_0, \dots, c_{k-1} \rangle$  with  $c_{k-1}(i) = b_{k-1}(i)$ . This says that agent  $i$  knows his own probability assignment.
- (B2) For all  $k > 1$ , the probability measure  $b_{k-1}(i)$  is the marginal of  $b_k(i)$  on  $X_{k-1}$ .

While we could in principle extend belief structures beyond the first  $\omega$  levels, just as we did in the case of knowledge structures, the result of [BE79, MZ85] assures us that there is no need; the first  $\omega$  levels determine the rest of the hierarchy.

---

<sup>2</sup>The assumption that  $S$  is a compact metric space is made in [TW88]. In [BD92] it is assumed that  $S$  is a complete separable metric space, while in [BE79] and [MZ85] it is simply assumed that  $S$  is compact. All these assumptions are trivially true if  $S$  is finite, which is often a reasonable assumption in practice.

### 3 Are knowledge structures adequate models of knowledge?

As shown in [FHV91], knowledge structures do not in general completely describe the agents' knowledge, and to fully capture this knowledge we need to extend the hierarchy into the transfinite ordinals. This should be contrasted with the situation for belief structures, which completely describe the agents' beliefs [BE79, MZ85]. To understand this difference better, the first question we want to examine here is when knowledge structures completely describe the agents' knowledge.

To answer this question, we first need to formalize it. If  $\lambda \geq \omega$  is an ordinal, then we say that a knowledge structure  $\mathbf{f}$  *characterizes the agents'  $\lambda$ -knowledge* if there is a unique extension of  $\mathbf{f} = \langle f_0, f_1, f_2, \dots \rangle$  to a  $(\lambda + 1)$ -world  $\langle f_0, f_1, f_2, \dots, f_\lambda \rangle$ . In particular,  $\mathbf{f}$  characterizes the agents'  $\omega$ -knowledge if the "next" level  $f_\omega$  is uniquely determined. We say that a knowledge structure  $\mathbf{f}$  (*completely*) *characterizes the agents' knowledge* if it characterizes the agents'  $\lambda$ -knowledge for every  $\lambda \geq \omega$ , that is, if all extensions of  $\mathbf{f}$  are determined. This definition captures the intuition that the first  $\omega$  levels determine the agents' knowledge. As we have already observed, the result of [BE79, MZ85] implies that all belief structures characterize the agents' beliefs in this sense.

We now provide a necessary and sufficient condition for a knowledge structure to characterize the agents' knowledge. First, we give a necessary and sufficient condition for a knowledge structure to characterize the agents'  $\omega$ -knowledge.

We say that a world  $\langle g_0, g_1, \dots, g_{k-1} \rangle \in f_k(i)$  is  *$i$ -uniquely extendible in  $\mathbf{f}$*  if there is a unique knowledge structure  $\mathbf{g} = \langle g_0, g_1, \dots, g_{k-1}, \dots \rangle$  such that  $\langle g_0, g_1, \dots, g_l \rangle \in f_{l+1}(i)$  for all  $l$ .

**Theorem 3.1:** *A knowledge structure  $\mathbf{f}$  characterizes the agents'  $\omega$ -knowledge iff for each agent  $i$  and each knowledge structure  $\mathbf{g} = \langle g_0, g_1, g_2, \dots \rangle$  different from  $\mathbf{f}$  such that  $\mathbf{f} \sim_i \mathbf{g}$ , there exists some  $r$  such that  $\langle g_0, g_1, \dots, g_r \rangle$  is  $i$ -uniquely extendible in  $\mathbf{f}$ .*

We might hope that if a knowledge structure characterizes the agents'  $\omega$ -knowledge, then it completely characterizes the agents' knowledge. Unfortunately, this is not the case. For example, agent 1 might consider it possible that agent 2's  $\omega$ -knowledge is not characterized. As the next result shows, a knowledge structure characterizes the agents' knowledge iff it is *common knowledge* that the first  $\omega$  levels characterizes the agents'  $\omega$ -knowledge. To make this precise, we need some more definitions.

Let  $\mathbf{f}$  and  $\mathbf{g}$  be knowledge structures. We say that  $\mathbf{g}$  is *reachable* from  $\mathbf{f}$  if there are knowledge structures  $\mathbf{h}_0, \dots, \mathbf{h}_r$  such that  $\mathbf{f} = \mathbf{h}_0$ ,  $\mathbf{g} = \mathbf{h}_r$ , and for all  $j < r$ , we have  $\mathbf{h}_j \sim_i \mathbf{h}_{j+1}$  for some agent  $i$ . By Theorem 2.1,  $\mathbf{f} \sim_i \mathbf{g}$  iff agent  $i$  considers every finite prefix of  $\mathbf{g}$  possible (according to  $\mathbf{f}$ ). Similarly,  $\mathbf{g}$  is reachable from  $\mathbf{f}$  iff, intuitively, according to  $\mathbf{f}$ , some agent considers it possible that some agent considers it possible  $\dots$  that some agent considers every finite prefix of  $\mathbf{g}$  possible. There is a close connection between reachability and common knowledge. For example, it can be shown that an event  $E \subseteq S$  is common knowledge in  $\mathbf{f}$  iff  $E$  holds at each knowledge structure reachable from  $\mathbf{f}$ . (See [Aum76, HM92] for analogous results in the context of Aumann structures, and [TW88] for an analogous result in the context of belief structures.)

**Theorem 3.2:** *A knowledge structure  $\mathbf{f}$  characterizes the agents' knowledge iff every knowledge structure reachable from  $\mathbf{f}$  characterizes the agents'  $\omega$ -knowledge.*

It is not hard to demonstrate knowledge structures that do and knowledge structures that do not characterize the agents' knowledge. For example, in [FHV91], there is a construction that, given



a  $k$ -world  $w$ , builds  $w^*$ , the *no-information extension* of  $w$ , a knowledge structure that, informally, is the knowledge structure where all each agent knows is what is already described by  $w$ . It is not hard to see from the construction there that  $w^*$  does not characterize the agents'  $\omega$ -knowledge. We shall see another example later (Example 3.6) where the knowledge structure does not characterize the agents'  $\omega$ -knowledge. An example of a knowledge structure that characterizes the agents' knowledge is one where the state of nature is common knowledge. This is a knowledge structure  $\mathbf{f} = \langle f_0, f_1, \dots \rangle$  where every  $f_k(i)$  is a singleton set. We leave to the reader the straightforward verification that such a knowledge structure characterizes the agents' knowledge.

As we noted, there exist knowledge structures that characterize the agents'  $\omega$ -knowledge, but do not completely characterize the agents' knowledge. Rather surprisingly, it turns out that if a knowledge structure  $\mathbf{f}$  characterizes the agents' knowledge through the first  $\omega \cdot 2 = \omega + \omega$  levels (that is, if  $\mathbf{f}$  characterizes the agents'  $(\omega + k)$ -knowledge for every natural number  $k$ ), then  $\mathbf{f}$  completely characterizes the agents' knowledge.

**Theorem 3.3:** *A knowledge structure characterizes the agents' knowledge iff it characterizes the agents' knowledge through the first  $\omega \cdot 2$  levels.*

To gain a better understanding of the issue of characterization of knowledge, we now consider a simpler sufficient condition on knowledge structures that guarantees characterization of the agents' knowledge. Let  $\mathbf{f}$  be a knowledge structure. A world *hereditarily appears in*  $\mathbf{f}$  if it is a prefix of a knowledge structure that is reachable from  $\mathbf{f}$ . Intuitively, a world  $w$  hereditarily appears in  $\mathbf{f}$  if some agent considers it possible that some agent considers it possible  $\dots$  that some agent considers  $w$  possible. Let  $k$  be a fixed natural number. We say that it is *common knowledge in*  $\mathbf{f}$  *how the state of nature determines the agents' knowledge* if whenever  $w = \langle g_0, \dots, g_r \rangle$  and  $w' = \langle g'_0, \dots, g'_r \rangle$  hereditarily appear in  $\mathbf{f}$ , and  $g_0 = g'_0$ , then  $w = w'$ . Using Theorem 2.1, it can be easily shown that it is common knowledge in  $\mathbf{f}$  how the state of nature determines the agents' knowledge precisely if whenever  $\mathbf{g}$  and  $\mathbf{g}'$  are reachable from  $\mathbf{f}$ , and the state of nature is the same in  $\mathbf{g}$  and  $\mathbf{g}'$ , then  $\mathbf{g} = \mathbf{g}'$ .

The next proposition gives us a simple sufficient condition on a knowledge structure that guarantees that it characterizes the agents' knowledge.

**Proposition 3.4:** *Let  $\mathbf{f}$  be a knowledge structure where it is common knowledge how the state of nature determines the agents' knowledge. Then  $\mathbf{f}$  characterizes the agents' knowledge.*

The interest in Proposition 3.4 comes from the fact that the way an agent determines what states are possible (or, in the case of belief structures, the way an agent determines how to assign probabilities) clearly ultimately depends on circumstances external to the agent, including perhaps what the agent has observed, the agent's upbringing, and a myriad of other influences. In many applications, the most natural way to model the state of nature will capture these external circumstances, and therefore it is common knowledge how the state of nature determines the agents' knowledge. The following simple example may clarify this.

**Example 3.5:** There are three agents, 1, 2, and 3. Consider a fact  $p$  such as "the price of IBM stock is over \$100". Suppose agents 1 and 3 discover whether or not  $p$  holds, and agent 2 does not. Agent 1 and 2 then start to communicate about  $p$  over an unreliable channel. First agent 1 tells agent 2 about  $p$ . If agent 2 receives the message, he sends an acknowledgement. If agent 1

receives the acknowledgement, he acknowledges the acknowledgement, and so on. If at any point a message is not received, there is no further communication. There is never any communication between agent 3 and the other two agents. We also assume that agent 3 has no idea how much time has transpired since agent 1 found out about  $p$ , so that, in particular, he has no upper bound on the number of rounds of messages that may have passed between agents 1 and 2. Thus, we can take  $S$  to consist of pairs of the form  $(q, k)$ , where  $q$  is either  $p$  or  $\bar{p}$ , and describes whether or not the fact  $p$  holds, and  $k$  describes how many messages have been exchanged between 1 and 2.

In this situation, it is common knowledge how the state of nature  $(q, k)$  determines the agents' knowledge. Intuitively, this is because once we know how many messages have been exchanged, we can determine each agent's knowledge. For example, suppose that the state of nature is  $(p, 2)$ , so that  $p$  holds and two messages have been exchanged (thus far) between 1 and 2 (i.e., 2 received 1's initial message, and 1 received 2's acknowledgement). Then at the first level, agent 1 considers the states  $(p, 2)$  and  $(p, 3)$  possible (since agent 1 does not know whether his acknowledgement to agent 2's last message was received by agent 2) and 2 considers the states  $(p, 1)$  and  $(p, 2)$  possible (since agent 2 does not know whether agent 1 received the last acknowledgement he sent). Agent 3 considers all worlds of the form  $(p, k)$ ,  $k \geq 0$  possible, since he knows  $p$  holds, but has no idea how many rounds of communication there have been. In the full paper, we show how we can continue this construction in a unique way, level by level. The key point here is that it is common knowledge how the state of nature  $(q, k)$  determines the agents' knowledge. Therefore, by Proposition 3.4, each knowledge structure that arises in this scenario characterizes the agents' knowledge.

Before leaving this example, let us consider exactly what knowledge the agents have in each of the knowledge structures that arise in this scenario. Let  $E$  be the event " $p$  holds" (so that  $E$  consists of all knowledge structures where the state of nature is of the form  $(p, k)$ ). It is easy to check that in (the knowledge structure that corresponds to) the state  $(p, 0)$ , the event  $K_1(E)$  holds but the event  $K_2(K_1(E))$  does not; in  $(p, 1)$ , the event  $K_2(K_1(E))$  holds but the event  $K_1(K_2(K_1(E)))$  does not; in  $(p, 2)$ , the event  $K_1(K_2(K_1(E)))$  holds but the event  $K_2(K_1(K_2(K_1(E))))$  does not; and so on. Thus, at no point does common knowledge of  $E$  ever hold between agents 1 and 2, where agents 1 and 2 are said to have common knowledge of  $E$  if both 1 and 2 know that both 1 and 2 know ... that  $E$  holds (cf. the discussion of the coordinated attack problem in [HM90]). Now consider agent 3. Then, informally, in every state agent 3 certainly knows that agents 1 and 2 do not have common knowledge of  $E$  (since they never attain common knowledge of  $E$  when communication is not guaranteed). He considers it possible, however, that agents 1 and 2 have arbitrarily deep knowledge of  $E$  (since agent 3 considers all the states  $(p, 0), (p, 1), (p, 2), \dots$  possible). ■

While in simple examples it does seem reasonable to include enough information in the state of nature so that it is common knowledge how the state of nature determines the agents' knowledge, in more complicated examples this becomes a serious modeling problem. For example, even if we accept that the sum total of an agent's upbringing, together with hereditary factors, completely determines the agent's knowledge, it is not clear that we want to include all this information in the state of nature when modeling, say, a simple game. Once we leave it out, however, the knowledge structure may no longer adequately model the agents' knowledge, as the following example shows.

**Example 3.6:** Suppose we consider the same situation as in Example 3.5, but change the description of the state of nature. Instead of the state of nature describing not only whether  $p$  is true, but also how many messages arrive, suppose we simply take the state of nature to describe whether or not  $p$  is true. Thus, there are only two states of nature,  $p$  and  $\bar{p}$ . Intuitively, in this case, in the

worlds we construct, the state of nature no longer determines the agents' knowledge. In fact, for no knowledge structure  $f$  that arises in such a scenario does  $f$  characterize the agents'  $\omega$ -knowledge, as we now show. Assume that  $f = \langle f_0, f_1, f_2, \dots \rangle$ . To capture the fact that agent 3 knows that agents 1 and 2 do not have common knowledge of  $E$ , it is necessary for  $f_\omega(3)$  to contain no  $\omega$ -world where agents 1 and 2 have common knowledge of  $E$ . This gives one possible extension of  $f$  (the one that properly captures the knowledge of the agents in our example). Since, however, agent 3 considers it possible that agents 1 and 2 have arbitrarily deep knowledge of  $E$ , it is not hard to see that there is another legal extension of  $f$ , where  $f_\omega(3)$  contains some  $\omega$ -world where agents 1 and 2 have common knowledge of  $E$ . ■

**Remark 3.7:** Proposition 3.4 can be strengthened in a number of straightforward ways. One is as follows: We say that it is *common knowledge how level  $k$  determines the agents' knowledge* if whenever  $w = \langle g_0, \dots, g_k, \dots, g_r \rangle$  and  $w' = \langle g'_0, \dots, g'_k, \dots, g'_r \rangle$  hereditarily appear in  $f$ , and their prefixes  $\langle g_0, \dots, g_k \rangle$  and  $\langle g'_0, \dots, g'_k \rangle$  are identical, then  $w = w'$ . Then Proposition 3.4 still holds when we replace "it is common knowledge how the state of nature determines the agents' knowledge" by "for some  $k$ , it is common knowledge how level  $k$  determines the agents' knowledge."

We can further strengthen Proposition 3.4 by further weakening the hypotheses: Let  $f$  be a knowledge structure, and let  $k$  be a fixed natural number. We say that *agent  $i$  knows that level  $k$  determines the agents' knowledge in  $f$*  if whenever  $g = \langle g_0, g_1, \dots \rangle$  and  $g' = \langle g'_0, g'_1, \dots \rangle$  are knowledge structures such that (a)  $f \sim_i g$ , (b)  $f \sim_i g'$ , and (c) the prefixes  $\langle g_0, \dots, g_k \rangle$  and  $\langle g'_0, \dots, g'_k \rangle$  are identical, then  $g = g'$ . Intuitively, this says that level  $k$  completely determines the knowledge structure, among those knowledge structures that agent  $i$  considers possible. Assume for now that  $f$  is a knowledge structure where for some  $k$ , each agent knows that level  $k$  determines the agents' knowledge. It turns out that this condition is *not* sufficient to guarantee that  $f$  completely characterizes the agents' knowledge, even if  $k = 0$ , that is, even if each agent knows that the state of nature determines the agents' knowledge. Nevertheless, we can show that this assumption (that for some  $k$ , each agent knows that level  $k$  determines the agents' knowledge) is sufficient to guarantee that the knowledge structure characterizes the agents'  $\omega$ -knowledge. We then see from Theorem 3.2 that if  $f$  is a knowledge structure where for some  $k$ , it is common knowledge that level  $k$  determines the agents' knowledge, then  $f$  characterizes the agents' knowledge.

Notice that the definition of common knowledge *that level  $k$  determines the agents' knowledge* is different from our earlier definition of common knowledge *how level  $k$  determines the agents' knowledge*. It is common knowledge in  $f$  that level  $k$  determines the agents' knowledge if in every knowledge structure  $g$  reachable from  $f$ , every agent knows that level  $k$  determines the agents' knowledge. It is possible, however, that there are two different knowledge structures  $g$  and  $g'$ , both reachable from  $f$ , that have the same prefix through level  $k$ . This cannot happen if it is common knowledge *how level  $k$  determines the agents' knowledge*. It is not hard to show (by making use of Theorem 2.1) that "common knowledge how" implies "common knowledge that". ■

How does the characterization of knowledge in knowledge structures relate to the characterization of beliefs in belief structures? To answer this question, we now consider an alternative way of capturing the intuition of when a knowledge structure captures the agents'  $\omega$ -knowledge. This time, we consider when it is the case that there is enough information in a knowledge structure to determine what other knowledge structures each agent considers possible.

Consider a knowledge structure  $f$ . If  $f$  characterizes the agents' knowledge, then the knowledge structures that agent  $i$  considers possible in  $f$  are precisely the knowledge structures in  $f^{\sim i} = \{g :$

$\mathbf{f} \sim_i \mathbf{g}$ . On the other hand, if agent  $i$  has *more* information than is described in  $\mathbf{f}$ , then he might consider only some proper subset of  $\mathbf{f}^{\sim i}$  possible. We would expect, however, that a knowledge structure  $\mathbf{f}$  would characterize agent  $i$ 's knowledge at each finite level. More precisely, we would expect that if  $w \in f_k(i)$  for some  $k$ , then there is a knowledge structure  $\mathbf{f}'$  that  $i$  considers possible such that  $w$  is a prefix of  $\mathbf{f}'$ . We say that a set  $\mathcal{P}$  of knowledge structures is a *coherent set of possibilities for agent  $i$  at  $\mathbf{f} = \langle f_0, f_1, f_2, \dots \rangle$*  if  $\mathbf{f} \in \mathcal{P}$  and

**P1.**  $\mathcal{P} \subseteq \mathbf{f}^{\sim i}$ .

**P2.** If  $w \in f_k(i)$  for some  $k > 0$ , then there is a knowledge structure  $\mathbf{f}' \in \mathcal{P}$  such that  $w$  is a prefix of  $\mathbf{f}'$ .

Thus, P1 says that the agent has at least as much information as is contained in  $\mathbf{f}$ , while P2 says that  $\mathbf{f}$  characterizes agent  $i$ 's knowledge at every finite level.

As expected, agent  $i$  always has at least one coherent set of possibilities at  $\mathbf{f}$ , namely  $\mathbf{f}^{\sim i}$ .

**Lemma 3.8:**  $\mathbf{f}^{\sim i}$  is a coherent set of possibilities for agent  $i$  at  $\mathbf{f}$ .

If there is only one coherent set of possibilities at  $\mathbf{f}$  for each agent  $i$ , then we might expect that  $\mathbf{f}$  characterizes the agents'  $\omega$ -knowledge. The following result shows that this is indeed the case.

**Theorem 3.9:** The knowledge structure  $\mathbf{f}$  characterizes the agents'  $\omega$ -knowledge iff there is only one coherent set of possibilities at  $\mathbf{f}$  for each agent  $i$ .

Thinking in terms of coherent sets of possibilities gives us some insight into why belief structures do characterize the agents' beliefs. We say that a set  $\mathcal{P}$  of knowledge structures is *limit closed* if a knowledge structure  $\mathbf{g} = \langle g_0, g_1, \dots \rangle$  is in  $\mathcal{P}$  whenever, for all  $k$ , there is a knowledge structure  $\mathbf{g}^k \in \mathcal{P}$  such that  $\langle g_0, \dots, g_k \rangle$  is a prefix of  $\mathbf{g}^k$ . Thus,  $\mathcal{P}$  is limit closed if, whenever all prefixes of a knowledge structure appear in  $\mathcal{P}$ , then the whole knowledge structure appears in  $\mathcal{P}$ .

It is easy to see that  $\mathbf{f}^{\sim i}$  is limit closed. A coherent set of possibilities need not, however, be limit closed in general. As the next result shows, if it is limit closed, then it must in fact be  $\mathbf{f}^{\sim i}$ .

**Proposition 3.10:**  $\mathcal{P}$  is a limit-closed coherent set of possibilities for agent  $i$  at  $\mathbf{f}$  iff  $\mathcal{P} = \mathbf{f}^{\sim i}$ .

Proposition 3.10 tells us that if *every* coherent set of possibilities for each agent  $i$  at  $\mathbf{f}$  is limit closed, then there is only one coherent set of possibilities, namely  $\mathbf{f}^{\sim i}$ . It then follows from Theorem 3.9 that  $\mathbf{f}$  characterizes the agents'  $\omega$ -knowledge. This result helps explain why belief structures characterize the agents' beliefs. Informally, we can view *countable additivity* as an analog of the limit-closure condition. Limit closure says that if all prefixes of a knowledge structure are possible, then so is the knowledge structure itself. Analogously, countable additivity says that the probabilities of the prefixes of a belief structure determine the probability of the belief structure. Thus, as a result of countable additivity, belief structures characterize the agents'  $\omega$ -beliefs. Since countable additivity is a built-in assumption in belief structures, it is common knowledge that belief structures characterize the agents'  $\omega$ -beliefs. Consequently, belief structures completely characterize the agents' beliefs.

The discussion so far focused on knowledge structures characterizing the agents' knowledge. We have seen, however, that, in general, knowledge structures do not characterize the agents'

knowledge. How serious a problem is this? How much of the agents' knowledge is captured by a knowledge structure? These questions are addressed in [FHV91] in a logic-theoretic framework. Roughly speaking, it is shown there that knowledge structures that agree on the first  $\omega$  levels cannot distinguish formulas involving only knowledge and common knowledge. We consider the same questions here in a logic-free manner. In particular, we try to understand the precise relationship between Aumann structures and knowledge structures.

One way to approach this issue is to consider an Aumann structure with  $\mathcal{F}(S)$  as its state space, in analogy with [BD92]. Given, however, that knowledge structures do not completely describe the agents' knowledge, it does not seem right to take the state space to be  $\mathcal{F}(S)$ . Instead, we consider a more general framework. Let us consider an Aumann structure with state space  $T$  such that every state in  $T$  is associated with a knowledge structure in  $\mathcal{F}(S)$ . Intuitively, we can think of the knowledge structure  $\mathfrak{f}_t$  associated with state  $t \in T$  as defining the agents' knowledge at state  $t$ , through the first  $\omega$  levels. We say that  $T$  is *sufficiently rich with respect to  $\mathcal{F}(S)$*  if every knowledge structure in  $\mathcal{F}(S)$  is associated with at least one state in  $T$ . We allow a knowledge structure to be associated with more than one state; since, as we have shown, knowledge structures do *not* in general completely characterize the agents' knowledge, there may be two distinct states of the world where the agents' knowledge through the first  $\omega$  levels may be identical. We say that a family of partitions  $\mathcal{K}_1, \dots, \mathcal{K}_n$  on  $T$  is *coherent* if, for every state  $t \in T$  and every agent  $i$ , the set of knowledge structures associated with the states in  $\mathcal{K}_i(t)$  form a coherent set of possibilities for agent  $i$ . As before, we can identify an event  $E \subseteq S$  with the set of all states in  $t \in T$  such that the state of nature in  $\mathfrak{f}_t$  is in  $E$ .

**Proposition 3.11:** *Suppose  $T$  is sufficiently rich with respect to  $\mathcal{F}(S)$ , and  $\mathcal{K}_1, \dots, \mathcal{K}_n$  and  $\mathcal{K}'_1, \dots, \mathcal{K}'_n$  are two coherent families of partitions on  $T$ . Consider the Aumann structures  $A_1 = (T, \mathcal{K}_1, \dots, \mathcal{K}_n)$  and  $A_2 = (T, \mathcal{K}'_1, \dots, \mathcal{K}'_n)$ . Assume  $E \subseteq S$  and  $t_0 \in T$ . Then  $E$  is common knowledge at the state  $t_0$  in the Aumann structure  $A_1$  iff  $E$  is common knowledge at the state  $t_0$  in the Aumann structure  $A_2$ .*

This result can be viewed as saying that knowledge structures do completely characterize the common knowledge that agents have regarding events in  $S$ . We are often interested, however, not just in common knowledge of events in  $S$ , but in common knowledge of more complicated events. For example, common knowledge of the partitions does not correspond to common knowledge of an event in  $S$ . As well, we are interested in what agents know as well as common knowledge. We can strengthen the previous result to include these cases.

Consider an Aumann structure of the form  $A = (T, \mathcal{K}_1, \dots, \mathcal{K}_n)$ . We can define the *ck-events over  $S$  in  $A$* , denoted  $ck_A(S)$ , as the result of starting with the events defined by subsets of  $S$ , and then closing off under complementation, finite intersection, and the knowledge and common knowledge operators.

**Proposition 3.12:** *Suppose  $T$  is sufficiently rich with respect to  $\mathcal{F}(S)$ , and  $\mathcal{K}_1, \dots, \mathcal{K}_n$  and  $\mathcal{K}'_1, \dots, \mathcal{K}'_n$  are two coherent families of partitions on  $T$ . Consider the Aumann structures  $A_1 = (T, \mathcal{K}_1, \dots, \mathcal{K}_n)$  and  $A_2 = (T, \mathcal{K}'_1, \dots, \mathcal{K}'_n)$ . Then  $ck_{A_1}(S) = ck_{A_2}(S)$ . Moreover, if  $E \in ck_{A_1}(S)$ , then  $E$  is known by agent  $i$  in  $A_1$  iff  $E$  is known by agent  $i$  in  $A_2$ , for  $i = 1, \dots, n$ , and  $E$  is common knowledge in  $A_1$  iff  $E$  is common knowledge in  $A_2$ .*

Thus, knowledge structures characterize the agents' knowledge even of the more complicated events described in the previous proposition. As was suggested in Example 3.6, the situation changes

when we consider also common knowledge among coalitions of agents. We can define a *coalition common knowledge operator*  $C_G$  in Aumann structures, for every coalition  $G$  of agents, along the same lines as we defined the common knowledge operator (the common knowledge operator is the special case where  $G$  is taken to be all the agents). Given an Aumann structure  $A_1$  as above, we can then define the *cck-events of  $S$  in  $A_1$* , denoted  $cck_{A_1}(S)$ , to be the result of closing off the sets of events also under the coalition common knowledge operators. As Example 3.6 suggest, Proposition 3.12 fails if we replace the ck-events by the cck-events. In this case it is necessary to carry the construction of the hierarchy into the transfinite ordinals. It turns out that if all we care about is the cck-events, then the construction needs to be carried out only up to the ordinal  $\omega^2$ . We simply state this result here, and leave details to the full paper.

Let  $\mathcal{F}^{\omega^2}(S)$  consist of all knowledge structures over  $S$  of length  $\omega^2$ . We can now consider an Aumann structure with state space  $T$  such that every state in  $T$  is associated with a knowledge structure in  $\mathcal{F}^{\omega^2}(S)$ . We can again define what it means for  $T$  to be sufficiently rich with respect to  $\mathcal{F}^{\omega^2}(S)$  and for  $\mathcal{K}_1, \dots, \mathcal{K}_n$  to be a coherent partition.

**Proposition 3.13:** *Suppose  $T$  is sufficiently rich with respect to  $\mathcal{F}^{\omega^2}(S)$ , and  $\mathcal{K}_1, \dots, \mathcal{K}_n$  and  $\mathcal{K}'_1, \dots, \mathcal{K}'_n$  are two coherent families of partitions on  $T$ . Consider the Aumann structures  $A_1 = (T, \mathcal{K}_1, \dots, \mathcal{K}_n)$  and  $A_2 = (T, \mathcal{K}'_1, \dots, \mathcal{K}'_n)$ . Then  $cck_{A_1}(S) = cck_{A_2}(S)$ . Moreover, if  $E \in cck_{A_1}(S)$ , and  $G$  is a group of agents, then  $E$  is joint knowledge in  $A_1$  for the group  $G$  iff  $E$  is joint knowledge in  $A_2$  for the group  $G$ , and  $E$  is known by agent  $i$  in  $A_1$  iff  $E$  is known by agent  $i$  in  $A_2$ , for  $i = 1, \dots, n$ .*

In the full paper, we also show that we need to consider structures of length  $\omega^2$  in order to get a result such as Proposition 3.13.

The focus so far has been on the issue of how much knowledge is captured by knowledge structures. As we observed, however, in the introduction, knowledge structures seem to also be deficient in another manner, since they capture only worlds that are commonly known to be possible, while omitting worlds that are merely conceivable (such as ones where Ron Fagin is President). Clearly, this deficiency is orthogonal to the issue of the length of the hierarchy. The answer to this argument is that the deficiency is a function of the definition of knowledge structures, and not of the hierarchical approach. In the full paper, we show how a generalization of knowledge structures can capture counterfactual information. We only provide a sketch here; details are left to the full paper.

The basic idea is to augment the definition of knowledge assignments. Assume inductively that extended  $k$ -worlds have been defined. Let  $U_k$  be the set of all extended  $k$ -worlds. A  *$k$ th-order extended knowledge assignment* is a function that associates with each agent  $i$  a set  $f_k^p(i) \subseteq U_k$  of “possible” extended  $k$ -worlds and a set  $f_k^c(i) \subseteq U_k$  of “conceivable” extended  $k$ -worlds such that  $f_k^p(i) \subseteq f_k^c(i)$ . Extended  $(k+1)$ -worlds are now defined as  $(k+1)$ -sequences of extended knowledge assignments that satisfy certain consistency conditions, which extend the consistency conditions that knowledge worlds are required to obey. Extended knowledge structures are now defined in analogy with knowledge structures. The manner in which extended knowledge structures enable counterfactual reasoning will be clarified in the full paper; intuitively, the conceivable worlds are used to model situations where Ron Fagin is President, which are known not to be possible.

## 4 Concluding Remarks

Our focus in this paper is on the expressive power of the hierarchical approach to modeling knowledge. Our results suggest a solution to the circularity problem of Aumann structures, since hierarchical structures can always serve as adequate models of the states in Aumann structures.

Beyond addressing the circularity problem, the hierarchical approach offers an attractive way to model knowledge: by a sequence of finite approximations. For example, it is shown in [FHV91] that this way of modeling facilitates the analysis of certain notions, such as the notion of “finite amount” of information.

It is interesting to note that the finiteness of the approximations is lost in the belief structures of [BE79, MZ85], since the second level already consists of uncountable probability spaces. We note that we can obtain an alternative approach to modeling probabilities using hierarchical structures. This approach is provably equivalent to the belief structures approach, but has the advantage that, if we start with a finite set  $S$  of states of nature, then at each level we have finite probability spaces.

**Acknowledgments:** We are grateful to Eddie Dekel for helpful discussions and comments.

## References

- [Aum76] R. J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4(6):1236–1239, 1976.
- [Aum89] R. J. Aumann. Notes on interactive epistemology. Cowles Foundation for Research in Economics working paper, 1989.
- [BD92] A. Brandenburger and E. Dekel. Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, to appear, 1992.
- [BE79] W. Böge and T. Eisele. On solutions of Bayesian games. *International Journal of Game Theory*, 8:193–215, 1979.
- [FHV91] R. Fagin, J. Y. Halpern, and M. Y. Vardi. A model-theoretic analysis of knowledge. *Journal of the ACM*, 91(2):382–428, 1991. A preliminary version appeared in *Proceedings of the 25th IEEE Symposium on Foundations of Computer Science*, 1984.
- [Gil88] I. Gilboa. Information and meta information. In M. Y. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 227–243, San Mateo, CA, 1988. Morgan Kaufmann, San Mateo, California.
- [Hal91] J. Y. Halpern. Reasoning about knowledge: a survey circa 1991. In A. Kent and J. G. Williams, editors, *Encyclopedia of Computer Science and Technology*. Marcel Dekker, 1991.
- [Har68] J. Harsanyi. Games with incomplete information played by ‘Bayesian’ players, parts I-III. *Management Science*, 14:159–182, 320–334, 486–502, 1968.
- [HC78] G. E. Hughes and M. J. Cresswell. *An Introduction to Modal Logic*. Methuen, London, 1978.

- [HM90] J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990. A preliminary version appeared in *Proceedings of the 3rd ACM Symposium on Principles of Distributed Computing*, 1984.
- [HM92] J. Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, to appear, 1992.
- [Kri59] S. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14, 1959.
- [MZ85] J. F. Mertens and S. Zamir. Formulation of Bayesian analysis for games of incomplete information. *International Journal of Game Theory*, 14(1):1–29, 1985.
- [TW88] T. C.-C. Tan and S. R. da C. Werlang. On Aumann's notion of common knowledge: an alternative approach. CARESS Working Paper 88-09, University of Pennsylvania, 1988.
- [Wer89] S. R. da C. Werlang. Common knowledge. In J. Eatwell, M. Milgate, and P. Newton, editors, *Game Theory*, pages 323–340. W. W. Norton, New York and London, 1989.