The Logic of Belief Persistency

Pierpaolo Battigalli and Giacomo Bonanno

Working Paper Series #95-18

Department of Economics
University of California
Davis, California 95616-8578

The Logic of Belief Persistency

Pierpaolo Battigalli* and Giacomo Bonanno

Working Paper Series No. 95-18
June 1995

* Department of Economics, Princeton University

Abstract

The interaction between knowledge and belief in a temporal context is analyzed. An axiomatic formulation and semantic characterization of the principle of belief persistency implied by the standard conditionalization rule are provided. This principle says that an individual does not change her mind unless new evidence forces her to do so. It is shown that if beliefs are conscious (or state-independent) and satisfy negative introspection then the principle of persistency of beliefs is characterized by the following axiom schema: the individual believes that ϕ at date t if and only if she believes at date t that she will believe that ϕ at date $t+1$.

1. Introduction

In the analysis of economic models with imperfect information the theorist ascribes two kinds of (non probabilistic) beliefs to the agents, which correspond to two nested epistemic levels:

(i) **"hard"** beliefs, given by the information that can be actually acquired in the economic interaction (usually described by means of information partitions), and

(ii) "soft" beliefs, representing what an agent is sure of in each specific situation (although the information actually acquired may be ***per*** se insufficient to obtain such certainty).

For example, in a discrete game in extensive form **"hard"** beliefs are given by the information sets, and "soft" beliefs are represented by the set of nodes in each information set having positive conditional (subjective) probability. It is normally understood that "hard" beliefs represent justified or veridical knowledge while "soft" beliefs might be arbitrary (as is the case in non equilibrium solution concepts such **as** rationalizability) and at most represent inferred knowledge that cannot be justified by observation alone. Following this interpretation, we adopt the convenient terminology of calling **"hard"** beliefs ***knowledge*** and "soft" beliefs simply ***beliefs.***

It may be argued that the distinction between knowledge and beliefs is not so clear-cut because any kind of epistemic state is necessarily hypothetical and, to some extent, unjustified.[1] Even mere observations are "theory laden" and to consider them as

---

[1] **In their textbook on game theory Osborne and Rubinstein (1994, p. 135) write: "In our view a model should attempt to capture the features of reality that the players *perceive;* it should not necessarily aim to describe the reality that an outside observer perceives, though obviously there are links between the two perceptions".**

hard facts is at best an abstraction.' Yet, there is a crucial difference between the two epistemic levels mentioned above. Although the actual information acquired in an economic interaction may be endogenously determined by the solution of the model, the information structure – semantically, what an agent knows at each state of the world – is exogenously given as part of the description of the model itself. On the other hand, the (soft) beliefs of an agent are endogenously determined by the solution of the model. They depend on the particular solution concept used and, for a fixed solution concept, on such fundamentals as the preferences of the agents. For example, for each information set of a (discrete) game the set of nodes with positive conditional probability depend on the equilibrium strategies.

A number of recent papers have shown that it is useful to analyze the epistemic aspects of decisions and social interaction using the tools of modal logic.[3] Modal logic provides a rich and flexible framework for a rigorous definition, discussion and characterization of epistemic assumptions. We are interested, in particular, in the analysis of the interaction between knowledge and belief in dynamic decision problems and games. In this paper we take a first step in this direction by providing an axiomatic formulation and a semantic chsuacterization of the minimal properties of beliefs implied by the standard conditionalization rule.

Consider an individual who in each period of time t may receive a new piece of information. In the standard semantic representation used by economists there is a set of

---

[2] The conjectural character of all knowledge is the central tenet of the epistemological approach broadly called *critical rationalism* (see, for example, the volume edited by Lakatos and Musgrave, 1970).

[3] For a list of references see the special issue of *Theory and Decision* on "Logic and the epistemic foundations of game theory" (1994, Vol. 37).

states and a sequence of **information partitions.**[4] If the individual assigns conditional (subjective) probabilities to the states, her beliefs correspond to the support of her conditional probability measure. Let S be the support at time t and let H be the new information set at time t+1. The standard conditionalization rule says that, if the intersection $S \cap H$ is non-empty, then $S \cap H$ must be the new support. Epistemically, S corresponds to the conjunction of all the propositions that the individual believes with certainty at t (that is, the conjectural theory of the individual at t) and H corresponds to the conjunction of all the **propositions** that she knows at t+1. The conditionalization rule implies that, as long as what the individual actually knows does not contradict her conjectural theory, she continues to believe in it and simply adds to it the propositions she has learned to be true. This captures an informal epistemic ***principle of persistency of beliefs: an*** individual does not change her mind unless new evidence forces her to do so. The conditionalization rule is illlustrated in Figure 1, where thick lines represent the information partition of the individual (her knowledge) and thin lines represent the support of her subjective conditional probability distribution (her beliefs). Thus, if the true state is, say, w_6, then at time t the individual knows (is informed) that the state is either w_4 or w_5 or w_6 and her subjective belief is that the true state is either w_4 or w_5 (she is "certain" that the true state is not w_6). At time t+1 she learns (is informed) that the true state is not w_4. By the **conditionalization** rule she must now attach probability 1 to state w_5.[5]

---

[4] Usually it is assumed that, as time progresses, the individual is given more information, that is her knowledge increases (she learns). This assumption is translated into the property that the information partition of the individual at time t+1 is a refinement of her information partition at time t.

[5] Similarly, if the true state is w_3, the individual knows, at date t, that the state is either w_1 or w_2 or w_3 and she attaches positive probability to all three. If at date t+1 she learns that the true state is not w_1, then she must attach positive probability to both w_2 and w_3.

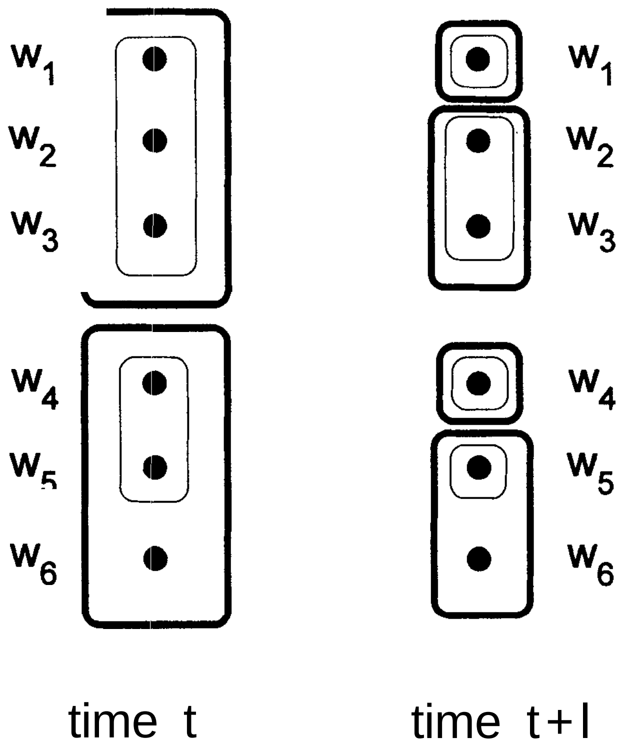

Figure 1

Despite the apparent simplicity of the principle of persistency of beliefs, an axiomatic formalization in terms of modal logic is not so straightforward. For example, the following axiom schema has been proposed (cf. Kraus and Lehmann, 1988):[6]

(PB) If, at date t, the individual believes that ϕ, then at date $t+1$ either she knows that ϕ is false or she still believes that ϕ.

[6] (PB) corresponds, in the framework of ow paper, to axiom schema (A21) in Kraus and Lehmann **(1988,** p. 107). The difference between their approach and ours is that we analyze situations where "the objective state of the world" does **not** change over time: the only thing that changes over time is the epistemic state of the individual. Thus time enters our analysis only through the knowledge and belief operators. In particular, unlike Kraus and Lehmann **(1988),** we do **not** have a time operator $\bigcirc$, where $\bigcirc\phi$ would be interpreted as "at the next date ϕ".

It easy to see that, in non trivial models, this axiom schema is unacceptable. Suppose that at date t the individual believes that both P and Q are true, and at date t+l she learns that either P or Q is false, but according to her new knowledge neither P nor Q can be ruled out as false. By (PIB), the individual must believe both P and Q at t+1, but this belief contradicts her knowledge. The problem with (PB) is that persistency of beliefs is postulated for every proposition believed by the individual. No problem would arise if persistency of beliefs were postulated for a single, specific proposition R. For example, there are solution concepts for dynamic games relying on the informal assumption that every player believes that the opponents are rational, as long as she does not observe behavior inconsistent with rationality (see, for example, Pearce, 1984; see also Kraus and Lehmann's, 1988, analysis of the "muddy children puzzle").

While postulating persistency of beliefs for specific propositions is an approach worth pursuing for the epistemic: analysis of dynamic economic models, it falls short of characterizing the basic notion of persistency implied by the conditionalization rule. Our previous discussion suggests that persistency of beliefs should be postulated for the composite proposition given by the conjunction of all the propositions believed by the individual, i.e. the theory of the individual. However, it is usually the case that the theory of the individual is given by an infinite set of propositions and an infinite conjunction of propositions is not a well-formed formula in the formal language of propositional (modal) logic. Thus we cannot formally use an axiom like "If T is the individual's theory at t and T is consistent with what the individual learns at t+1, then the individual believes T at t+1." We solve this problem by showing that, given other standard axioms and inference rules for knowledge and belief, the rule of conditionalization is characterized by the following axiom schema:[7]

(PB') The individual believes that ϕ at date t if and only if she believes at date t that she will believe that ϕ at date $t+1$.[8]

The formal language that we put forward in Section 2 is the one that comes closest to the dynamic models developed in the information economics literature. In particular, we restrict our analysis to situations where the objective state of the world does not change over time, that is, the truth value of the atomic propositions (which provide a factual description of the world) is constant over time. The only thing that varies with time is the epistemic state of the individual, that is, what the individual knows and believes about the world. Thus time enters our analysis only through the knowledge and belief operators.

The paper is organized as follows. In Section 2 we develop the formal analysis. Section 3 contains an extended discussion of the main result. Section 4 contains a conclusion and a discussion of related literature.

---

[7] In particular, we require that beliefs be state-independent or conscious (if the individual believes that ϕ then she knows that she believes that ϕ) and that they satisfy negative introspection (if the individual does not believe that ϕ then she believes that she does not believe that ϕ).

[8] The reader might have noticed the formal similarity between this axiom and the law of iterated expectations.

2. Characterization of belief persistency

Let $T \subseteq \mathbb{N}$ (where $\mathbb{N}$ is the set of non-negative integers). We consider a logic with two modal operators for every $t \in T$: B_t and K_t. The intended interpretation of $B_t\phi$ is "at time t the individual **believes** that ϕ" and the interpretation of $K_t\phi$ is "at time t the individual knows[9] that ϕ".[10] The alphabet of the language consists of: (1) a finite or countable set $\Pi = \{\pi_1, \pi_2, ...\}$ of ***sentence letters*** (representing atomic propositions), **(2)** a set $T \subseteq \mathbb{N}$ of ***dates*** satisfying the property that if $t \in \mathbb{N}$ and $t+1 \in T$ then $t \in T$, **(3)** the ***connectives*** $\neg$ (for "not"), v (for "or"), and, for every $t \in T$, B_t and K_t, (4) the ***bracket symbols*** (and). A ***word*** is a finite string of elements of the alphabet. The set Φ of ***formulae*** is the subset of the set of words defined recursively as follows:

(i) for every sentence letter IT, $(\pi) \in$ O,

(ii) if $\phi \in \Phi$ then $(\neg\phi) \in 0$, and, for every $t \in T$, $(B_t\phi) \in \Phi$ and $(K_t\phi) \in 0$,

(iii) if $\phi, \psi \in \Phi$ then $(\phi \vee \psi) \in \Phi$.

As is customary, we shall often omit the outermost brackets (e.g. we shall write $\phi \vee \psi$ instead of $(\phi \vee \psi)$). Furthermore, we shall use the following metalinguistic abbreviations: ϕ A for $\neg(\neg\phi$ v $\neg\psi)$ (the symbol A stands for "and") and $\phi \rightarrow \psi$ for $(\neg\phi)$ v ψ (the symbol $\rightarrow$ stands for "if...then..").

---

[9] As explained in the introduction, the distinction between K_t and B_t ought to be thought of as a distinction between "hard" beliefs (not necessarily knowledge) and "soft" beliefs. In particular, our main result does not require the Axiom of Truth for K_t: $K_t\phi \rightarrow \phi$.

[10] As explained in the introduction, our aim is to analyze situations where the only thing that changes over time is the epistemic state of the individual: the factual statements that describe the world do not change with time. Thus time enters our analysis only through the knowledge and belief operators.

We denote by $\mathbf{K}^{\mathbf{time}}$ the ***system*** or ***calculus*** specified by the following axiom schemata and rules of inference:

(1) All the tautologies (that is, a suitable axiomatization of propositional calculus),

(2) the schema K (cf. Chellas, 1980):

$K_t(\phi \to \psi) \to (K_t\phi \to K_t\psi)$, for every $t \in T$,

$B_t(\phi \to \psi) \to (B_t\phi \to B_t\psi)$, for every $t \in T$,

(3) the rule of inference *Modus Ponens*:

$$\text{MP} \qquad \frac{\phi\,,\ \phi \to \psi}{\psi}$$

(4) the rule of inference ***Necessitation:***

$$\frac{\phi}{K_t\phi} \qquad \text{for every } t \in T,$$

$$\frac{\phi}{B_t\phi} \qquad \text{for every } t \in T.$$

We now turn to the semantics. ***A standard frame*** is a tuple

$$\langle W,\ T,\ \{\mathcal{K}_t\}_{t\in T},\ \{\mathcal{B}_t\}_{t\in T}\rangle$$

where

(1) W is a set of ***worlds*** or states, whose elements are denoted by u, v, w ...

(2) $T \subseteq \mathbb{N}$ is such that if $t \in \mathbb{N}$ and $t+1 \in T$ then $t \in T$.

(3) For every $t \in T$, $\mathcal{K}_t$ is a binary relation on W (intuitively $v\mathcal{K}_t w$ means that, at time t, if the true state is v then the individual considers w possible, i.e. cannot rule out w).

(4) For every $t \in T$, $\mathcal{B}_t$ is a **binary** relation on W (intuitively $v\mathcal{B}_t w$ means that, at time t, if the true state is v then the individual considers w likely, that is, attaches positive probability to w).

A standard model is a tuple $\mathcal{M} = (W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T}, f)$ where $(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T})$ is a standard frame and $f : \Pi \to 2^W$ (2^W denotes the set of subsets of W). For every propositional variable π, $f(\pi)$ is the set of worlds at which π is true. We say that $\mathcal{M}$ is ***based on*** the frame $(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T})$.

Given a formula ϕ and a standard model $\mathcal{M} = (W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T}, f)$, the ***truth set*** of ϕ in $\mathcal{M}$, denoted by $\|\phi\|^{\mathcal{M}}$ is defined recursively as follows:

(1) If $\phi = (n)$ where π is a sentence letter, then $\|\phi\|^{m} = f(\pi)$,

(2) $\|\neg\phi\|^{\mathcal{M}} = W - \|\phi\|^{\mathcal{M}}$ (that is, $\|\neg\phi\|^{\mathcal{M}}$ is the complement of $\|\phi\|^{\mathcal{M}}$)

(3) $\|\phi \vee \psi\|^{m} = \|\phi\|^{\mathcal{M}} \cup \|\psi\|^{\mathcal{M}}$,

(4) For all $t \in T$

$\|K_t\phi\|^{\mathcal{M}} = \{ u \in W :$ for all v such that $u\mathcal{K}_t v$, $v \in \|\phi\|^{\mathcal{M}} \}$, and

$\|B_t\phi\|^{\mathcal{M}} - \{ u\,s\,W :$ for all v such that $u\mathcal{B}_t v$, $v \in \|\phi\|^{\mathcal{M}} \}$.

If $v \in \|\phi\|^{\mathcal{M}}$ we say that ϕ ***is true at world v in model*** $\mathcal{M}$. ***An*** alternative notation for $v \in \|\phi\|^{\mathcal{M}}$ is $\models_v^{\mathcal{M}} \phi$ and. an alternative notation for $v \notin \|\phi\|^{m}$ is $\not\models_v^{\mathcal{M}} \phi$. **A** formula ϕ is ***valid in model*** $\mathcal{M}$ if and only if $\models_v^{\mathcal{M}} \phi$ for all $v \in W$.

Let P be a property of the relations $\mathcal{B}_t$ and/or the relations $\mathcal{K}_t$ and o be an axiom schema We say that o is ***characterized by property P*** if: (i) every instance of o is valid in every model based on a frame that satisfies property P, and (ii) given a frame that violates property P, there exist a model $\mathcal{M}$ based on it and an instance ϕ of σ that is not valid in $\mathcal{M}$. For example, it is well known (see Chellas, 1980) that axiom schema (known as negative introspection) $\neg B_t\phi \rightarrow B_t\neg B_t\phi$ (respectively, $\neg K_t\phi \rightarrow K_t\neg K_t\phi$) is characterized by the property that $\mathcal{B}_t$ (respectively, $\mathcal{K}_t$) is euclidean,[11] axiom schema (known as positive introspection) $B_t\phi \rightarrow B_tB_t\phi$ (respectively, $K_t\phi \rightarrow K_tK_t\phi$) is characterized by the property that $\mathcal{B}_t$ (respectively, $\mathcal{K}_t$) is transitive, axiom schema $K_t\phi \rightarrow \phi$ (known as veridicality) is characterized by the property that $\mathcal{K}_t$ is reflexive, etc.

We are interested in the system obtained by adding the following axiom schemata to the system $\mathbf{K}^{\mathbf{time}}$.

(A1) $\quad B_tB_{t+1}\phi \rightarrow B_t\phi$

(A2) $\quad B_t\phi \rightarrow B_tB_{t+1}\phi$

(A3) $\quad B_t\phi \rightarrow K_tB_t\phi$

(A1) says that if at date it the individual believes that she will believe that ϕ at date t+1, then she must believe that ϕ at date t. (A2) says the converse: if she believes that ϕ at date t, then she must also believe, at date t, that she will believe that ϕ at date t+1. (A3) says that beliefs are conscious: if the individual believes that ϕ then she knows that she believes this.

---

[11] Recall that $\mathcal{B}_t$ is euclidean if $u\mathcal{B}_tv$ and $u\mathcal{B}_tw$ implies $v\mathcal{B}_tw$.

PROPOSITION 1. The following characterization holds

(1) Axiom schema (A1) is characterized by the following property

(R1) $\forall$ u,v$\in$ W, $\forall$ t$\in$ T, if $u\mathcal{B}_t v$ and $(t+1)\in T$ then $\exists$ w$\in$ W such that $u\mathcal{B}_t w$ and $w\mathcal{B}_{t+1} v$.

(2) Axiom schema (A2) is characterized by the following property

(R2) $\forall$ u,v,w$\in$ W, $\forall$ t$\in$ T, if $u\mathcal{B}_t v$ and $v\mathcal{B}_{t+1} w$ then $u\mathcal{B}_t w$.

(3) Axiom schema (A3) is characterized by the following property

(R3) $\forall$ u,v,w$\in$ W, $\forall$ t$\in$ T, if $u\mathcal{K}_t v$ and $v\mathcal{B}_t w$ then $u\mathcal{B}_t w$.

Proposition 1 can be seen as an application of Theorem 4.3 (c and e) in van der Hoek (1993, p. 183).[12] For the reader's convenience, and because van der Hoek does not provide a complete proof, we give the proof of Proposition 1 in the appendix. (Note that van der Hoek's analysis has a completely different focus from ours: he investigates the ***atemporal*** relationship between knowledge and belief, in particular, the maximal "consciousness" conditions compatible with the non-collapse of belief into knowledge.)

Our objective in this paper is to provide an axiomatic characterization of the notion of belief persistency corresponding to the rule of conditionalization. Semantically, the notion that beliefs are ***persistent*** – that is, an individual keeps on believing her previous theory until she knows it is false -- is captured by property P which is the conjunction of the following two properties:

(P1) $\forall$ u,v$\in$ W, $\forall$ t$\in$ T, if $u\mathcal{B}_t v$ and $u\mathcal{K}_{t+1} v$ then $u\mathcal{B}_{t+1} v$,

[12] We are grateful to Joe Halpern for bringing this paper to our attention.

(P2) $\forall u \in W, \forall t \in T$, if $\exists v \in W$ such that $u\mathcal{B}_{t+1}v$ and not $u\mathcal{B}_t v$ then $\forall w \in W$ if $u\mathcal{B}_t w$ then not $u\mathcal{K}_{t+1}w$.

Property (P1) says that if v is belief-accessible from u at time t and is knowledge-accessible from u at time t+1, then it is also belief-accessible at time t+1. Thus property (P1) ***rules out arbitrary contractions*** of the belief set (cf. Figure 1). Property (P2) says that if, at time t+1, v is belief-accessible from u despite the fact that it was not at time t, then it must be the case that every w which was belief-accessible from u at time t is not knowledge-accessible from u at time t+1. Thus (P2) ***rules out arbitrary expansions*** of the belief set (cf. Figure 1).

PROPOSITION 2. If property (R3) is satisfied and, for all $t \in T$, $\mathcal{B}_t$ is euclidean, then the conjunction of properties (R1) and (R2) implies the conjunction of properties (P1) and (P2).

Proof: First we prove (P1). Fix arbitrary u and v such that $u\mathcal{B}_t v$ and $u\mathcal{K}_{t+1}v$. We need to show that $u\mathcal{B}_{t+1}v$. By (R1) there exists a w such that $u\mathcal{B}_t w$ and $w\mathcal{B}_{t+1}v$. Since $\mathcal{B}_{t+1}$ is euclidean, $v\mathcal{B}_{t+1}v$. Since: $u\mathcal{K}_{t+1}v$ and $v\mathcal{B}_{t+1}v$, by (R3) it follows that $u\mathcal{B}_{t+1}v$.

Next we prove (P2). Note that (P2) can be written as (is equivalent to) $\forall u,v,w \in W$, $u\mathcal{B}_{t+1}v \;\&\; u\mathcal{K}_{t+1}w \;\&\; u\mathcal{B}_t w \Rightarrow u\mathcal{B}_t v$. Fix arbitrary u,v and w such that $u\mathcal{B}_{t+1}v$, $u\mathcal{K}_{t+1}w$ and $u\mathcal{B}_t w$. By (P1) [which was proved above], since $u\mathcal{B}_t w$ and $u\mathcal{K}_{t+1}w$, it follows that $u\mathcal{B}_{t+1}w$. By euclicleanness of $\mathcal{B}_{t+1}$, since $u\mathcal{B}_{t+1}v$ and $u\mathcal{B}_{t+1}w$, we have that $w\mathcal{B}_{t+1}v$. This, together with $u\mathcal{B}_t w$, yields, by (R2), $u\mathcal{B}_t v$. ■

We postpone until the next section a discussion of what is needed in order to prove a partial converse of Proposition 2.

The following proposition, together with Proposition 2, identifies a system that provides an axiomatization of the: notion of persistency of beliefs (for a further **discussion** see the next **section**).[13]

PROPOSITION 3. Let Σ be the system obtained by adding to $\mathbf{K}^{\text{time}}$ the following axiom schemata: for every $t \in T$, $\neg B_t\phi \rightarrow B_t\neg B_t\phi$ (negative introspection of beliefs), (A1), (A2) and (A3). Then Σ is sound and complete with respect to the class of models where: (1) $\forall\, t \in T$, $\mathcal{B}_t$ is euclidean, (2) properties **(R1),** (R2) and (R3) are satisfied.

Proposition 3, again, can be seen as an application of Theorem 4.3 (2) in van der Hoek (1993, p. 183). For the reader's convenience, and because van der Hoek does not provide a complete proof, we give the proof of Proposition 3 in the appendix.

COROLLARY 1. The system Σ **axiomatizes** the notion of belief persistency.

Proof: It follows from Proposition 3 and Proposition 2. ■

Notice that, under many respects, Σ is a weaker system than the one normally used in applications. In particular, modeling knowledge by means of information partitions implies assuming **veridicality** ($K_t\phi \rightarrow \phi$), and negative introspection ($\neg K_t\phi \rightarrow K_t\neg K_t\phi$), for the **knowledge** operator, which are not assumed in Proposition 3. Indeed, no assumptions concerning the knowledge operator are made in Proposition 3 (aside from axiom (A3) which concerns the relation between knowledge and belief),

---

[13] Recall that a logic is sound and complete with respect to a class *C* of models if: (i) every theorem of the logic (that is, every formula that can be derived from the axioms by means of the rules of inference) is valid in every model in *C*, and (ii) every formula that is valid in every model in *C* is a theorem of the logic.

although one would probably want to require at least KD45 (or weak S5) for knowledge.[14] Similarly, modeling beliefs with (the support of conditional) probability distributions implies (see Halpern, 1991) assuming not only negative introspection, but also consistency ($B_t\phi \rightarrow \neg B_t\neg\phi$) and positive introspection ($B_t\phi \rightarrow B_tB_t\phi$), for the belief operator, which are not assumed in Proposition 3. For a further discussion of this point see the next section. The following example, illustrated in Figure 2, shows the principle of belief persistency applied to a situation where both knowledge and belief satisfy the logic of KD45 (semantically, $\mathcal{K}_t$ and $\mathcal{B}_t$, for t = 0, 1, are serial,[15] transitive and euclidean), but knowledge is not veridical, that is, the Truth Axiom does not hold for K_t (semantically, the relation $\mathcal{K}_t$ is not reflexive).

EXAMPLE 1. T = {0, 1}, W = {u, v, w, x}, $\mathcal{K}_0$ = {(u,v), (u,w), (u,x), (v,v), (v,w), (v,x), (w,v), (w,w), (w,x), (x,v), (x,w), (x,x)}, $\mathcal{K}_1$ = {(u,v), (u,w), (v,v), (v,w), (w,v), (w,w), (x,x)}, 4 = { (u,w), (u,x), (v,w), (v,x), (w,w), (w,x), (x,w), (x,x)}, $\mathcal{B}_1$ = { (u,w), (v,w), (w,w), (x,x)}. In Figure 2 the relation $\mathcal{K}_t$ is represented by thick arrows and thick shapes, while the relation $\mathcal{B}_t$ is represented by thin arrows and thin shapes (if a set S of worlds is enclosed in a thick shape, then the relation $\mathcal{K}_t$ restricted to this set is universal, that is, $y\mathcal{K}_tz$ for all $y,z \in$ S; similarly for thin shapes and the relation $\mathcal{B}_t$). Thus if the true state is u, then at date 0 the individual considers v, w and x possible

---

[14] **In Chellas (1980) KD45 is the system where the knowledge operator satisfies the following axiom schemata:**

K. $K_t(\phi \rightarrow \psi) \rightarrow (K_t\phi \rightarrow K_t\psi)$

D. $K_t\phi \rightarrow \neg K_t\neg\phi$ **(consistency)**

4. $K_t\phi \rightarrow K_tK_t\phi$ **(positive introspection)**

5. $\neg K_t\phi \rightarrow K_t\neg K_t\phi$ **(negative introspection).**

It the Truth **Axiom ($K_t\phi \rightarrow \phi$) is added then the corresponding system is called S5 or KT5.**

[15] **A relation R is serial if for every u there is a v such that uRv.**

(i.e. cannot rule out any of them), but attaches positive probability only to w and x. At date 1 the individual learns that the true state is not x and she now attaches probability 1 to w. Let π_1 be a proposition whose truth set is {v, w, x) and π_2 a proposition whose truth set is {w,x}. Then at date 0 and at state u, the individual knows (and believes) that π_1 and does not know but believes that π_2. At date 1 and state u the individual still (knows and believes that π_1 and) believes that π_2. At both dates the individual is wrong in her knowledge and belief.

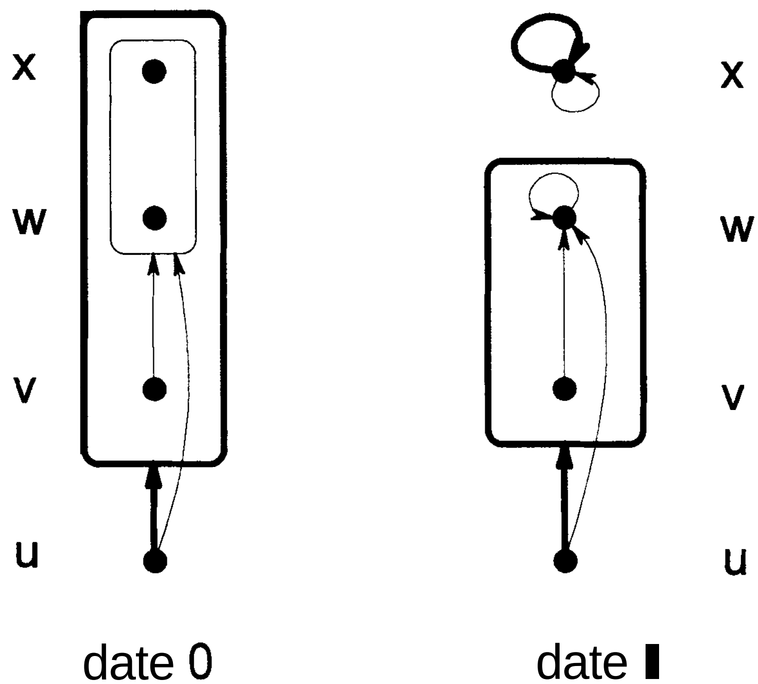

Figure 2

3. Discussion

It was remarked after Corollary 1 that, although Proposition 3 does not require any assumptions about the knowledge operator and only negative introspection for the belief operator (as well as the axiom that beliefs are conscious), reasonable axiomatizations of knowledge (or "hard beliefs") would require at least consistency and positive and negative introspection, that is, at least the logic of **KD45** (or weak S5: cf. footnote 14). Semantically, this translates into the requirement that $\mathcal{K}_t$ be serial, transitive and euclidean. Furthermore, if beliefs (that is, "soft" beliefs) at a world u are represented by the support of a probability distribution over the set of nodes that are knowledge-accessible from u, then $\mathcal{B}_t$ would also satisfy **KD45** (see Halpern, 1991). Moreover, for beliefs to be based on knowledge, it is also necessary to postulate that the individual believes everything that she knows:

$$\text{(A4)} \qquad K_t\phi \rightarrow B_t\phi$$

The following lemma is proved in the appendix.

LEMMA **1**. Axiom schema (A4) is characterized by the following property

$$\text{(R4)} \quad \forall\, u,v \in W,\ \forall\, t \in T,\ \text{ if } u\mathcal{B}_t v \text{ then } u\mathcal{K}_t v.$$

Furthermore, the standard canonical model. (see the appendix) of a system that contains (A4) satisfies property (R4).

From now on we shall restrict attention to systems where both the knowledge and the belief'operators satisfy the logic of **KD45** and, furthermore, axiom schemata (A3) (consciousness of beliefs) and (A4) (what is known is believed) are postulated.

Restricting attention to such systems, does the converse of Proposition 2 hold? The answer is negative, as the following example shows.

EXAMPLE 2. Let T = {1,2}, W = {u,v}, $\mathcal{K}_0 = \mathcal{B}_0$ = {(u,v),(v,v)} and $\mathcal{K}_1$ = 23, = {(u,u),(v,u)}. This model, which is illustrated in Figure 3, satisfies the following properties: (1) for every t = 0,1, $\mathcal{K}_t$ and $\mathcal{B}_t$ are serial, transitive and euclidean, (2) properties (R3), (R4), (P1) and (P2) are satisfied. Yet both (R1) and (R2) are violated.

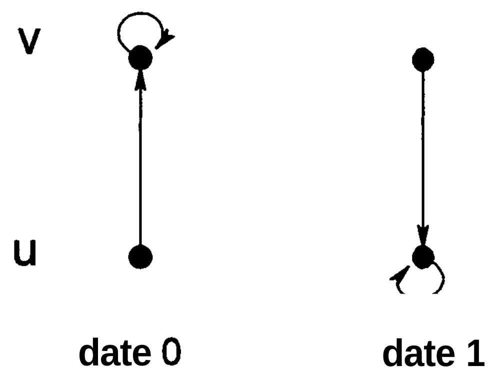

Figure 3

At each date t = 0,1, $\mathcal{K}_t = \mathcal{B}_t$ is represented by arrows.

However, if the Truth Axiom is added for the knowledge operator: $K_t\phi \rightarrow \phi$, that is, if knowledge satisfies the logic of KT5 (or S5: see footnote 14) then the conjunction of (P1) and (P2) becomes equivalent to the conjunction of (R1) and (R2), as the following proposition shows.

PROPOSITION 4. Suppose that properties (R3) and (R4) are satisfied and, for every $t \in T$, $\mathcal{B}_t$ is euclidean and $\mathcal{K}_t$ is reflexive. Then the conjunction of (P1) and (P2) is equivalent to the conjunction of (R1) and (R2).

Proof: By Proposition 2 it is enough to prove that (P1) & (P2) implies (R1) & (R2). In fact, we will show the stronger result that (P1) implies (R1) and (P2) implies (R2).

(P1) $\Rightarrow$ (R1): Let $u,v \in W$ be such that $u\mathcal{B}_t v$. We want to show that there exists a $w \in W$ such that $u\mathcal{B}_t w$ and $w\mathcal{B}_{t+1} v$. Choose w = v. Then we only have to show that $v\mathcal{B}_{t+1} v$. By reflexivity of $\mathcal{K}_{t+1}$, $v\mathcal{K}_{t+1} v$. By euclideanness of $\mathcal{B}_t$, $v\mathcal{B}_t v$. By (P1), since $v\mathcal{B}_t v$ and $v\mathcal{K}_{t+1} v$, $v\mathcal{B}_{t+1} v$.

(P2) $\Rightarrow$ (R2): Let $u,v,w \in W$ be such that $u\mathcal{B}_t v$ and $v\mathcal{B}_{t+1} w$. We need to show that $u\mathcal{B}_t w$. By reflexivity of $\mathcal{K}_{t+1}$, $v\mathcal{K}_{t+1} v$. By euclideanness of $\mathcal{B}_t$, $v\mathcal{B}_t v$. Thus we have:

$$v\mathcal{B}_{t+1} w \text{ and } v\mathcal{K}_{t+1} v \text{ and } v\mathcal{B}_t v.$$

By (P2) this implies $v\mathcal{B}_t w$. By (R4), since $u\mathcal{B}_t v$, $u\mathcal{K}_t v$. By (R3), since $u\mathcal{K}_t v$ and $v\mathcal{B}_t w$, it follows that $u\mathcal{B}_t w$. ■

To conclude our discussion, we shall consider a fifth, and last, axiom schema:

$$\text{(A5)} \qquad K_t\phi \rightarrow K_{t+1}\phi.$$

This axiom captures the notion of perfect memory or recall: if the individual knows that ϕ at date t then she will know that ϕ at every future date. In the case where knowledge is represented by information partitions, (A5) corresponds to the semantic assumption that the information partition of the individual at time t+1 is a refinement of her information partition at time t. The following lemma is proved in the appendix

LEMMA 2. Axiom schema (A5) is characterized by the following property

$$\text{(R5)} \quad \forall u,v \in W, \forall t \in T, \text{ if } u\mathcal{K}_{t+1} v \text{ then } u\mathcal{K}_t v.$$

Furthermore, the standard canonical model of a system that contains (A5) satisfies property (R5).

Axiom (A5) plays no role in our results. Thus our axiomatization of the conditionalization rule applies also to situations where memory is lacking, as shown in the example of Figure 4 below (where, as before, thick lines denote the information partitions that represent knowledge, and thin lines denote the supports of the conditional probability distributions that represent beliefs).

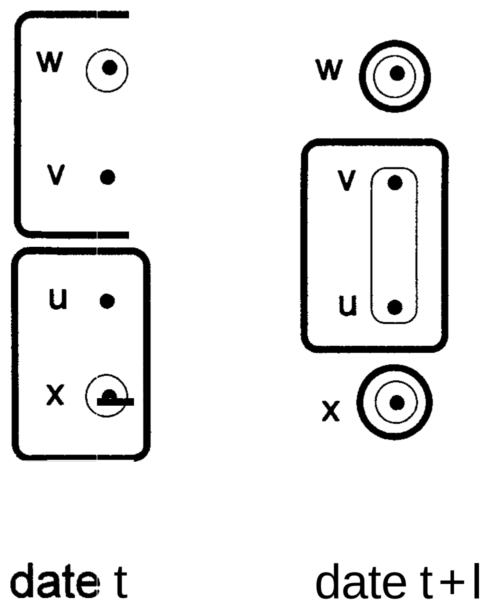

Figure 4

4. Related literature

We conclude with a review of related literature. The *atemporal* relationship between knowledge and belief was first analyzed in Kraus and Lehmann (1 988). In particular, the atemporal version of our axioms (A3) (consciousness of beliefs) and (A4) (what is known is also believed) can be found there. Kraus and Lehmann postulated the full S5 logic for knowledge, consistency $(B\phi \rightarrow \neg B\neg\phi)$ for beliefs, and (A3) and (A4) for the interaction between knowledge and beliefs. They showed that positive and negative introspection for the belief operator are theorems of this system. Kraus and Lehmann also considered a multi-agent logic with operators for common knowledge and common belief. In the last part of the paper the authors considered the possibility of extending the logic to include a time operator. In particular, they addressed the question of how to characterize the notion of persistency of beliefs: "if person i believes something, he will keep on believing it until he knows it is false" (1988, p. 107). They listed, and briefly discussed, a number of possible axioms (we mentioned, and criticized, one of them in the introduction) and concluded by saying that "An open problem is: find a natural family of models for which the systems considered above are complete".

One property of the system considered by Kraus and Lehmann is that if one adds the axiom schema $B\phi \rightarrow BK+$ then knowledge and belief become identical, that is, one obtains the theorem $B\phi \leftrightarrow K\phi$. This point is taken up by van der Hoek (1993) in an extensive analysis of the causes of this "problem" and of a similar system that allows one to introduce the axiom $B\phi \rightarrow BK\phi$ without obtaining a collapse of belief into knowledge.

An extensive analysis of' knowledge in a temporal context can be found in Halpern and Vardi (1989). Among the issues considered are: whether or not the individual[16] forgets, whether or not she learns, whether or not time is synchronous, and whether or not there is a unique initial state in the system. The objective of their paper is to characterize the complexity of the validity problem for all the logics considered.

Somewhat related is also Scherl and Levesque (1995). The authors use situation calculus to model actions and their effects on the world. Axioms are used to specify the prerequisites of actions as well as their effects, that is, the fluents that they change. The analysis centers on knowledge-:producing actions, that is, actions whose effects are to change a state of knowledge. Knowledge is modeled as veridical: reflexivity of the accessibility relation for knowledge turns out to be crucial for their results. An interesting

---

[16] **The authors are actually interested in modeling knowledge and time for distributive systems and therefore talk about the knowledge of a processor, rather than an individual.**

aspect of Scherl and Levesque's analysis is that memory emerges as a side-effect: if something is known in a certain situation, it remains known at successor situations, unless something relevant has changed.

One more paper which is relevant to the issues considered here is Halpern (1991), which studies the relation between knowledge and certainty, where a fact is ***known*** if it is true at all worlds an individual considers possible and ***certain*** if it holds with probability 1. Halpem shows that if one assumes one fixed probability assignment (such an assumption would correspond, in our framework, to axiom (A3)) then the logic KD45 provides a complete axiomatization for reasoning about certainty. However, Halpem does not deal with the issue of the evolution of knowledge and belief over time.

Some of the papers reviewed above deal with the apparently more general case where there are $n \geq 1$ individuals, whereas we have restricted attention to the case of one individual. It should be clear, however, that our results apply also to the multi-agent case (the only modification required in the statement of the results and in the proofs is the attachment of a superscript i to the epistemic operators and the accessibility relations, where the index i ranges over the set of agents).

Appendix

Proof of Proposition 1. Proof of (1). (A). Let $\mathcal{M}$ be a model that satisfies property (R1). Fix arbitrary $u \in W$, $t \in T$ and an arbitrary formula ϕ. Suppose that $\models_u^{\mathcal{M}} B_t B_{t+1}\phi$. We want to show that $\models_u^{\mathcal{M}} B_t\phi$, that is, that for all v such that $u\mathcal{B}_t v$, $\models_v^{\mathcal{M}} \phi$. Fix an arbitrary v such that $u\mathcal{B}_t v$. By the assumed property, there exists a w such that $u\mathcal{B}_t w$ and $w\mathcal{B}_{t+1} v$. Since $u\mathcal{B}_t w$ and $\models_u^{\mathcal{M}} B_t B_{t+1}\phi$, $\models_w^{\mathcal{M}} B_{t+1}\phi$ and, since $w\mathcal{B}_{t+1} v$, $\models_v^{\mathcal{M}} \phi$.

(B). Let $\left(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T}\right)$ be a frame that violates property (R1). Then there exist $t \in T$ and $u,v \in W$ such that $u\mathcal{B}_t v$ and, for all $w \in W$, if $u\mathcal{B}_t w$, then not $w\mathcal{B}_{t+1} v$. Let π be a propositional variable and $\mathcal{M}$ a model where the truth set of π is the set of worlds that can be reached from u in two steps, first with $\mathcal{B}_t$ and then with $\mathcal{B}_{t+1}$, that is, $f(\pi) = \{z \in W : \text{for some } x \in W,\ u\mathcal{B}_t x \text{ and } x\mathcal{B}_{t+1} z\}$. Then $v \notin f(\pi)$. Hence (since $u\mathcal{B}_t v$) $\not\models_u^{\mathcal{M}} B_t\pi$. On the other hand, by definition of $f(\pi)$, $\models_u^{\mathcal{M}} B_t B_{t+1}\pi$.

Proof of (2). (A). Let $\mathcal{M}$ be a model that satisfies property (R2). Fix arbitrary u s W, $t \in T$ and an arbitrary formula ϕ. Suppose that $\models_u^{\mathcal{M}} B_t\phi$. We want to show that $\models_u^{\mathcal{M}} B_t B_{t+1}\phi$. Fix arbitrary v and w such that $u\mathcal{B}_t v$ and $v\mathcal{B}_{t+1} w$. By the assumed property, $u\mathcal{B}_t w$. Hence, since $\models_u^{\mathcal{M}} B_t\phi$, $\models_w^{\mathcal{M}} \phi$.

(B). Let $\left(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T}\right\rangle$ be a frame that violates property (R2). Then there exist $t \in T$ and $u,v,w \in W$ such that $u\mathcal{B}_t v$ and $v\mathcal{B}_{t+1} w$ and not $u\mathcal{B}_t w$. Let π be a propositional variable and $\mathcal{M}$ a model where $f(\pi) = \{z \in W : u\mathcal{B}_t z\}$. Thus $\models_u^{\mathcal{M}} B_t\pi$. On the other hand, since $w \notin f(\pi)$, $\not\models_u^{\mathcal{M}} B_t B_{t+1}\pi$.

Proof of (3). (A). Let $\mathcal{M}$ be a model that satisfies property (R3). Fix arbitrary u s W, $t \in T$ and an arbitrary formula ϕ. Suppose that $\models_u^{\mathcal{M}} B_t\phi$. We want to show that

$\models^{\mathcal{M}}_{u} K_tB_t\phi$. Fix arbitrary v and w such that $u\mathcal{K}_tv$ and $v\mathcal{B}_tw$. By the assumed property, $u\mathcal{B}_tw$. Hence, since $\models^{\mathcal{M}}_{u} B_t\phi$, $\models^{\mathcal{M}}_{w} \phi$.

(B). Let $(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T})$ be a frame that violates property (R3). Then there exist $t\in T$ and $u,v,w\in W$ such that $u\mathcal{K}_tv$ and $v\mathcal{B}_tw$ and not $u\mathcal{B}_tw$. Let π be a propositional variable and $\mathcal{M}$ a model where $f(\pi) = \{z\in W : u\mathcal{B}_tz\}$. Thus $\models^{\mathcal{M}}_{u} B_t\pi$. On the other hand, since $w\notin f(\pi)$, $\not\models^{\mathcal{M}}_{u} K_tB_t\pi$. ■

Proof of Proposition 3. (1) follows from a standard soundness and completeness theorem for modal logic (see Chellas, 1980). That (A1), (A2) and (A3) are valid in this class of models follows from Proposition 1. Thus we only need to prove completeness. We proceed in the usual way. Let $\mathcal{M} = (W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T}, f : \Pi \to 2^W)$ be the standard canonical model for Σ. That is, W is the set of maximal consistent sets of formulae, $u\mathcal{K}_tv$ iff $\{\phi : K_t\phi\in u\} \subseteq v$ and $u\mathcal{B}_tv$ iff $\{\phi : B_t\phi\in u\} \subseteq v$. Furthermore, for every sentence letter π, $f(\pi) = \{w\in W : \pi\in w\}$. To prove completeness it is enough to show that the canonical model satisfies properties (R1), (R2) and (R3) (cf. Chellas, 1980).

Proof of (R1). Choose arbitrary $t\in T$ and $u,v\in W$ such that $u\mathcal{B}_tv$, that is, $\{\phi : B_t\phi\in u\} \subseteq v$. We want to show that there exists a $w\in W$ such that $\{\phi : B_t\phi\in u\} \subseteq w$ and $\{\phi : B_{t+1}\phi\in w\} \subseteq v$. By Theorem 4.29 in Chellas (1980, p. 158), $\{\phi : B_{t+1}\phi\in w\} \subseteq v$ if and only if $\{\neg B_{t+1}\neg\psi : \psi\in v\} \subseteq w$. Thus we want to find a $w\in W$ such that $\Gamma \cup A \subseteq w$, where $\Gamma = \{\phi : B_t\phi\in u\}$ and $A = \{\neg B_{t+1}\neg\psi : \psi\in v\}$. By Lindenbaum's lemma, this is equivalent to showing that $\Gamma \cup A$ is consistent. Suppose it is not consistent. Then there exist $\phi_1, ..., \phi_n\in\Gamma$ and $\neg B_{t+1}\neg\psi_1, ..., \neg B_{t+1}\neg\psi_m\in A$ (with $n \geq 0$, $m\geq 0$ and $n+m \geq 1$) such that $\neg(\phi_1 \wedge ... \wedge \phi_n \wedge \neg B_{t+1}\neg\psi_1 \wedge ... \wedge \neg B_{t+1}\neg\psi_m)$ is a theorem of Σ. [Note that it must be $m \geq 1$ because, otherwise, we would have that $\neg(\phi_1 \wedge ... \wedge \phi_n)$

is a theorem of Σ, contradicting the assumption that, for every $i = 1, ..., n$, $\phi_i \in v$ and v is a maximal consistent set of formulae.] By propositional logic this is equivalent to $(\phi_1 \wedge ... \wedge \phi_n) \rightarrow (B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m)$ [in the case where $n = 0$, we would have that $(B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m)$ is a theorem of Σ]. By the rule of inference RK (see Chellas, 1980, p. 121) for B_t, it follows that $(B_t\phi_1 \wedge ... \wedge B_t\phi_n) \rightarrow B_t(B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m)$ is a theorem of Σ: [in the case where $n = 0$, by the rule of necessitation for B_t, we would have that $B_t(B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m)$, which is (i) below, is a theorem of Σ]. Hence it belongs to u. Since, for every $i = 1, ..., n$, $B_t\phi_i \in u$ (because $\phi_i \in \Gamma$), it follows that $(B_t\phi_1 \wedge ... \wedge B_t\phi_n) \in u$ and therefore

$$B_t(B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m) \in u \qquad \text{(i)}$$

Since $(B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m) \rightarrow B_{t+1}(\neg\psi_1 \vee ... \vee \neg\psi_m)$ is a theorem of every normal system (see Chellas, 1980, p. 123), it belongs to u. Hence by the rule of inference RM (see Chellas,, 1980, p.114) the following formula belongs to u: $B_t(B_{t+1}\neg\psi_1 \vee ... \vee B_{t+1}\neg\psi_m) \rightarrow B_t B_{t+1}(\neg\psi_1 \vee ... \vee \neg\psi_m)$. It follows from (i) that $B_t B_{t+1}(\neg\psi_1 \vee ... \vee \neg\psi_m) \in u$. Since Σ contains axiom schema (A1), the following formula is in u: $B_t B_{t+1}(\neg\psi_1 \vee ... \vee \neg\psi_m) \rightarrow B_t(\neg\psi_1 \vee ... \vee \neg\psi_m)$. Hence $B_t(\neg\psi_1 \vee ... \vee \neg\psi_m)$ belongs to u. Since $u\mathcal{B}_t v$, it follows that $(\neg\psi_1 \vee ... \vee \neg\psi_m) \in v$. By propositional logic, $(\neg\psi_1 \vee ... \vee \neg\psi_m)$ is equivalent to $\neg(\psi_1 \wedge ... \wedge \psi_m)$. Hence

$$\neg(\psi_1 \wedge ... \wedge \psi_m) \in v \qquad \text{(ii).}$$

On the other hand, for every $j = 1, ..., m$, $\psi_j \in v$ (since $\neg B_{t+1}\neg\psi_j \in \Delta$). Thus

$$(\psi_1 \wedge ... \wedge \psi_m) \in v \qquad \text{(iii).}$$

But (ii) and (iii) together imply that v is inconsistent, contradicting the assumption that $v \in W$, that is, that v is a maximal consistent set of formulae.

Proof of (R2). Fix **arbitrary** $t \in T$ and $u,v,w \in W$ such that $u\mathcal{B}_t v$ and $v\mathcal{B}_{t+1} w$. Choose an arbitrary formula ψ such that $B_t\psi \in u$. We need to show that $\psi \in w$. Since, by (A2), $(B_t\psi \to B_tB_{t+1}\psi)$ is a theorem of Σ, $(B_t\psi \to B_tB_{t+1}\psi) \in u$. Hence $B_tB_{t+1}\psi \in u$. Since $u\mathcal{B}_t v$, $B_{t+1}\psi \in v$ and since $v\mathcal{B}_{t+1} w$, $\psi \in w$.

Proof of (R3). Fix **arbitrary** $t \in T$ and $u,v,w \in W$ such that $u\mathcal{K}_t v$ and $v\mathcal{B}_t w$. Choose an arbitrary formula ψ such that $B_t\psi \in u$. We need to show that $\psi \in w$. Since, by (A3), $(B_t\psi \to K_tB_t\psi)$ is a theorem of Σ, $(B_t\psi \to K_tB_t\psi) \in u$. Hence $K_tB_t\psi \in u$. Since $u\mathcal{K}_t v$, $B_t\psi \in v$ and since $v\mathcal{B}_t w$, $\psi \in w$. ■

Proof of Lemma 1. (A). Let $\mathcal{M}$ be a model that satisfies property **(R4)**. Fix arbitrary $u \in W$, $t \in T$ and an arbitrary formula ϕ. Suppose that $\models^{\mathcal{M}}_{u} K_t\phi$. Choose an arbitrary v such that $u\mathcal{B}_t v$. Then, by the assumed property, $u\mathcal{K}_t v$, hence $\models^{\mathcal{M}}_{v} \phi$.

(B). Let $(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T})$ be a **frame** that violates property **(R4)**. Then there exist $t \in T$ and $u,v \in W$ such that $u\mathcal{B}_t v$ and not $u\mathcal{K}_t v$. Let π be a propositional variable and $\mathcal{M}$ a model where $f(\pi) = W - \{v\}$. Then $\models^{\mathcal{M}}_{u} K_t\pi$. On the other hand, since $v \notin f(\pi)$, $\not\models^{\mathcal{M}}_{u} B_t\pi$.

Now fix a system that contains axiom schema (A4) and consider the corresponding standard **canonical** model. We want to show that it satisfies property **(R4)**. Fix arbitrary $t \in T$ and $u,v \in W$ such that $u\mathcal{B}_t v$. Choose an arbitrary formula ψ such that $K_t\psi \in u$. We need to show that $\psi \in v$. Since, by **(A4)**, $(K_t\psi \to B_t\psi)$ is a theorem of the system, $(K_t\psi \to B_t\psi) \in u$. Hence $B_t\psi \in u$. Since $u\mathcal{B}_t v$, $\psi \in v$. ■

Proof of Lemma 2. (A). Let $\mathcal{M}$ be a model that satisfies property (R5). Fix arbitrary u s W, $t \in T$ and an arbitrary formula ϕ. Suppose that $\models^{\mathcal{M}}_{u} K_t\phi$. Choose an arbitrary v such that $u\mathcal{K}_{t+1}v$. By the assumed property, $u\mathcal{K}_t v$, hence $\models^{\mathcal{M}}_{v} \phi$.

(B). Let $\left(W, T, \{\mathcal{K}_t\}_{t\in T}, \{\mathcal{B}_t\}_{t\in T}\right)$ be a frame that violates property (R5). Then there exist $t\in T$ and $u,v\in W$ such that $u\mathcal{K}_{t+1}v$ and not $u\mathcal{K}_t v$. Let π be a propositional variable and $\mathcal{M}$ a model where $f(\pi) = \{w\in W : u\mathcal{K}_t w\}$. Then $\models^{\mathcal{M}}_{u} K_t\pi$. On the other hand, since $v\notin f(\pi)$, $\not\models^{\mathcal{M}}_{u} K_{t+1}\pi$.

Now fix a system that contains axiom schema (A5) and consider the corresponding standard canonical model. We want to show that it satisfies property (R5). Fix arbitrary $t\in T$ and $u,v\in W$ such that $u\mathcal{K}_{t+1}v$. Choose an arbitrary formula ψ such that $K_t\psi\in u$. We need to show that $\psi\in v$. Since, by (AS), $(K_t\psi \rightarrow K_{t+1}\psi)$ is a theorem of the system, $(K_t\psi \rightarrow K_{t+1}\psi)\in u$. Hence $K_{t+1}\psi\in u$. Since $u\mathcal{K}_{t+1}v$, $\psi\in v$. ■

References

Chellas, B. (1980), Modal logic: an introduction, Cambridge University Press, Cambridge (U.K.).

Halpem, J. (1991), The relationship between knowledge, belief and certainty, Annals of Mathematics and *Artificial* Intelligence, 4, 301-322.

Halpern, J. and M. Vardi (1989), The complexity of reasoning about knowledge and time. I. Lower bounds, *Journal* of Computer and Systems Sciences, 38, 195-237.

van der Hoek, W. (1993), Systems for knowledge and belief, Journal of Logic and Computation, 3, 173-195.

Kraus, S. and D. Lehmann (1988), Knowledge, belief and time, Theoretical Computer Science, 58, 155-174.

Lakatos, I. and A. Musgrave (Eds.) (1970), Criticism and the growth of knowledge, Cambridge UK, Cambridge University Press.

Osborne, M. and A. Rubinstein (1994), A course in game theory, M.I.T. Press, Cambridge (MA).

Pearce, D. (1984), Rationalizable strategic behavior and the problem of perfection, *Econometrica*, 52, 1029-1050.

Scherl, R. and H. Levesque (1995), The frame problem and knowledge-producing actions, mimeo, Department of Computer Science, University of Toronto.