ELSEVIER Mathematical Social Sciences 30 (1995) 127–153

mathematical social sciences

Belief closure: A semantics of common knowledge for modal propositional logic

Luc Lismont[a], Philippe Mongin[b,c,*]

[a]GREQAM, École des Hautes Études en Sciences Sociales, 2 rue de la Charité, F-13002 Marseille, France

[b]Centre National de la Recherche Scientifique, France

[c]CORE, Université Catholique de Louvain, 34 voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium

Received August 1994; revised January 1995

Abstract

The paper investigates the relations between *iterate* and *fixed-point* accounts of common belief and common knowledge, using the formal tools of epistemic modal logic. Its main logical contribution is to introduce and axiomatize the following (fixed-point) notion of common belief. We first define a proposition to be *belief-closed* if everybody believes it in every world where it is true. We then define a proposition to be *common belief* in a world if it is implied by a belief-closed proposition that everybody believes in that world. Using the belief closure semantics of common belief, the paper proves soundness and completeness theorems for modal logics of varying strength. The weakest system involves a monotonicity assumption on individual belief; the strongest system is based on S5. Axiomatizations of *common knowledge* are secured by adding the truth axiom to any system. The paper also discusses anticipations of the belief closure semantics in the economic and game-theoretic literatures.

Keywords: Belief closure; Common belief; Common knowledge; Modal logic

1. Introduction and overview

Nontechnically, a proposition is common knowledge if it is true, if every individual knows it, if every individual knows that every individual knows it, and

* Correspondence to: P. Mongin, CORE, voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. Tel: (32)-10-474328 or 4315; Fax: (32)-10-474301; E-mail: mongin@core.ucl.ac.be.

SSDI 0165-4896(95)00785-7

so on ad infinitum. The significance of the common knowledge concept has come to be recognized by game theorists, mathematical economists, Artificial Intelligence as well as computer scientists, and philosophical logicians. In the hands of these researchers, it has led to numerous separate developments. The time seems to be ripe to investigate the analogies and disanalogies between the various approaches to common knowledge. The authors' recent survey paper (Lismont and Mongin, 1994a) brought to the attention of game theorists and mathematical economists a sample of the work recently pursued by AI scientists and logicians on this topic. This paper contrasted the informal (set-theoretic) method employed by the former with the logical approach to knowledge and belief, which involves the use of a *formal language* along with set-theoretic methods.

The present paper has a comparative purpose of a different, more technical sort. Most of the currently available definitions of common knowledge can be classified as being of either the *iterate* or the *circular* (or *fixed-point*) kinds, to use Barwise's (1989) expressions. The authors' primary target here is to relate these alternative views of common knowledge to each other.

Definitions of the former group just elaborate on the nontechnical one: they formalize the infinite regress of shared knowledge some way or another. Definitions of the latter group are not so close to the basic intuition as are the former, but they have proved to be both easier to handle and conceptually richer. One should expect of a fixed-point definition that it will imply some variant of the iterate definition, while also implying the following property: common knowledge of a proposition is essentially equivalent to everybody's knowledge of the common knowledge of that proposition, everybody's knowledge of everybody's knowledge of the common knowledge of that proposition, and so forth. This further property is clearly of the circular type. It points to the fact that, contrary to shared knowledge, common knowledge *itself* does not give rise to any infinite regress. In other words, one intuitively feels that infinite hierarchies of shared knowledge do not normally collapse into first-order knowledge; this is why one needs a common knowledge concept. But one also feels that common knowledge is all one needs. Infinite hierarchies of common knowledge should collapse into first-order common knowledge.

The distinction between iterate and fixed-point accounts cuts across the boundaries of several fields of inquiry. As will be argued below, it can be found in both the game theorists' and logicians' work, but emerges much more clearly from the logician's work. This is why we shall pursue most of our analysis within the confines of *epistemic modal logic*: it proves to be the most elegant and flexible framework for the present comparative purpose. However, the various connections of our results with the recent game-theoretic contributions will be emphasized in due course.

To start with a now classic view of common knowledge, consider Aumann's (1976). He first defines it in terms of the meet of the individuals' information partitions, and then explains that this definition can be rephrased into more intuitive terms, using the notion of a 'reachable' state of the world. Very roughly speaking, definition 1 is of the circular kind, and definition 2 of the iterate kind. However, in view of the immediate mathematical equivalence between definitions 1 and 2, this interpretative comment strikes one as far-fetched. Clearly, Aumann's framework is not rich enough to suggest interesting differences between the iterate and fixed-point accounts of common knowledge. More will be said later on the game theorists' further contributions on this score. But Aumann's article is by and large representative of what can be expected from employing only straightforward set-theoretic methods.

The situation is not at all the same in epistemic modal logic. There, the duality of iterate versus fixed-point approaches normally interacts with the distinction between syntax (i.e. the formal language) and semantics (i.e. the informal meta-language). 'Normally' in the last sentence is intended to exclude infinitary logics: we shall here retain the standard practice of permitting only finite conjunctions in the syntax.[1] Given this constraint, the syntactical account of common knowledge is *bound* to be circular. We shall indeed exploit a definition of the C operator by means of a *fixed-point axiom* (FP) and a *rule of induction* (RI) that is currently in favour in AI and logic (Fagin et al., 1991; Halpern and Moses, 1990, 1992; Lismont, 1993; Lismont and Mongin, 1991, 1994a,b). The linguistic constraint is relaxed when it comes to semantics: both iterate and circular accounts then become available. As a matter of fact, epistemic modal logic has opened up the following three alternative possibilities.

(1) Fagin, Halpern, Moses and Vardi (forthcoming) (henceforth referred to as FHMV) semantically define common knowledge of a statement in terms of the natural iteration: everybody knows that statement; everybody knows that everybody knows it, and so on ad infinitum. They have proved or stated a number of axiomatization theorems based on this semantics, using either the standard concept of a Kripke structure (e.g. Halpern and Moses, 1992) or the novel concept of a knowledge structure that they show to be closely related to the latter (e.g. Fagin et al., 1991). Barring these technical facts, the general interpretation of their results is as follows: they clarify the relations holding between the circular syntactical definition of common knowledge and the iterate definition chosen for the semantics.

(2) At the other extreme, as it were, the present paper will give a try to the following, completely fixed-point semantics. Roughly speaking, a proposition will be defined to be common knowledge if it is implied by a proposition that, for one, happens to be known by everybody, for another, has the following special

[1] For an infinitary logic approach to common knowledge, see Kaneko and Nagashima (1991, 1993). These authors criticize finitary axiomatizations of the standard type. The infinitary logic approach is also adopted in Heifetz (1994a).

property of *belief closure*: it is believed (known) at every state of the world where it is true. Abstracting from the logical context, this definition is by and large analogous to earlier ones stated by the economist Milgrom (1981), and the game theorists Mertens and Zamir (1985), and Monderer and Samet (1989). A proper comparison will be made in due course. It suffices now to mention the common theme: there is a privileged connection between common knowledge propositions and those special propositions (referred to as belief-closed), *which cannot hold without everybody believing them*. To investigate the various facets of the belief closure semantics is an important technical step in the strategy of this paper. Using the concepts of a neighborhood and a Kripke semantics in succession, it will prove several soundness and completeness theorems. Informally, these results should be viewed as stating the relations that hold between the circular syntactical definition of common knowledge and the also circular definition now chosen for the semantics.

(3) The last semantics to be considered does not fit easily within the iterate/fixed-point classification. Essentially, it amounts to interpreting the syntactical common knowledge operator in terms of the existing interpretations of the syntactical individual knowledge operators. In the context of Kripke structures, this method leads to a semantic definition of the well-known 'transitive closure' type (see, for example, Halpern and Moses, 1992, or FHMV, forthcoming). It is only in the more general context of neighborhood structures that method (3) fully manifests its technical potentialities. Lismont (1993) introduced a neighborhood semantics definition of the common knowledge operator that is phrased in terms of the neighborhood counterparts of the individual knowledge operators. This construction leads to several soundness and completeness theorems; it has a fixed point as well as an iterate component.

To sum up the state of the art, the modal logic of common knowledge typically involves a circular syntax, whereas the semantics is diverse. The syntax/semantics duality creates a rich framework of analysis in which the question of iterate versus circular accounts of common knowledge can be addressed.

The paper is organized as follows. Section 2 introduces the syntactical definitions. The formal language is of the usual modal propositional type; its unary operators are intended to represent individual, shared, and common belief or knowledge, respectively. The axioms are selected among those most widely discussed in epistemic logic and AI. Some purely syntactical facts about common belief and common knowledge are collected in Proposition 1. Section 3 introduces the belief closure semantics in the context of neighborhood structures. It then proves a soundness and completeness theorem for a weak system consisting of a fixed-point axiom (FP), a rule of induction (RI), and a monotonicity rule imposed on both the individual and common belief operators (Theorem 2). Section 4 defines the belief closure semantics in the case of Kripke structures and then derives soundness and completeness results for systems consisting of (FP), (RI), the monotonicity rule of common belief, and various systems of individual belief: Theorem 3 deals with Kripke's minimal system, while Theorem 4 covers several stronger systems, including the classic S4 and S5. Once the logical groundwork has been completed, it becomes possible to move to the major question of this paper. Section 5 addresses it by mutually comparing approaches (1), (2) and (3). Proposition 5 and Theorem 6 summarize the various implications holding between them. It turns out that complete equivalence prevails in the Kripkean context, but breaks down in the more general neighborhood context. This final section comments on these results and stresses the connections between our belief closure semantics and relevant anticipations from the economic and game-theoretic literatures.[2]

2. Syntactical definitions and facts

The formal language and axiom systems discussed in this paper derive their special features from the fact that there are belief (knowledge) operators B_a, one for each individual or 'agent' a, and, even more importantly, a specific operator C to render 'it is common belief (knowledge) that'.

Before proceeding further, we should clarify our underlying distinction between knowledge and belief. Following a widespread view, what is known must be true, whereas what is believed may be either true or false. Contemporary epistemic logic drastically simplifies the analysis of knowledge versus belief by taking *only* this difference into account. That is to say, it analyzes knowledge as a particular case of belief, i.e. the particular case in which the believed statement or proposition is true. Despite its well-recognized philosophical defects, we shall follow the epistemic logician's practice and (at least from now on) conform our terminology with it.

More precisely, we shall take the intended meaning of the B_a to be 'a believes that' in the general case, and to be 'a knows that' only exceptionally, when the so-called truth axiom holds ('what is believed is true'). Consistently, we shall interpret C to mean 'it is common belief that' in the general case, and 'it is common knowledge that' only in the special case where the truth axiom holds. In view of this terminological distinction, most of the work currently pursued under the heading of 'common knowledge' is, in fact, concerned with common belief. The present paper is no exception to this state of affairs.

[2] Some prior knowledge of the basic tools of epistemic modal logic, especially Kripke and neighborhood structures, might facilitate the appreciation of the present results. The usual recommended introduction to modal logic at large is Chellas (1980). FHMV's (forthcoming) textbook is a specifically epistemic treatment of various modal logic systems. The reader might also consult the companion paper to this one (Lismont and Mongin, 1994a). However, the present paper provides the relevant definitions and is therefore essentially self-contained.

The formal language of our systems is constructed from the following building blocks:

(i) a set PV of propositional variables of any cardinality;

(ii) the logical connectives $\wedge$, $\vee$, $\neg$, $\rightarrow$, $\leftrightarrow$; and

(iii) the unary operators: $(B_a)_{a \in A}$, C and E.

The requirement that the set of individuals A be finite is crucial for this paper. The added *shared-belief* operator E is introduced for convenience (its intended meaning is 'everybody believes that'). Let Φ denote the set of well-formed formulas constructed from these components following the obvious closure rules. The letters φ and ψ will refer to typical elements of Φ.

Our 'minimal' system MC_A is made out of any axiomatization of the propositional calculus (by means of suitable tautologies) and the following modal rules and axiom schemata:

$$(\mathrm{RM}_a) \quad \frac{\varphi \rightarrow \psi}{B_a\varphi \rightarrow B_a\psi}, \quad \text{for any } a \in A;$$

$$(\mathrm{Def}.E) \quad E\varphi \leftrightarrow \bigwedge_{a \in A} B_a\varphi;$$

$$(\mathrm{RM}_C) \quad \frac{\varphi \rightarrow \psi}{C\varphi \rightarrow C\psi};$$

$$(\mathrm{FP}) \quad C\varphi \rightarrow E(C\varphi \wedge \varphi);$$

$$(\mathrm{RI}) \quad \frac{\varphi \rightarrow E\varphi}{E\varphi \rightarrow C\varphi}.$$

The *monotonicity rules* (RM_a) mean that the logic defined by the above system and the agents' logic share a strong common component. However, MC_A is compatible with the following two limiting cases: an agent does not believe anything at all; he believes in a contradiction. These two limiting cases disappear from the stronger systems than MC_A, which also include, for any $a \in A$,

$$(\mathrm{N}_a) \quad B_a\top,$$

$$(\mathrm{P}_a) \quad \neg B_a\bot,$$

where $\top$ and $\bot$ stand for any propositional theorem and contradiction, respectively. Even when MC_a is enriched with (P_a), the agent might entertain a belief in φ and a belief in $\neg\varphi$ 'without drawing the consequences'. This case is excluded from the systems that also include (C_a):

$$(\mathrm{C}_a) \quad B_a\varphi \wedge B_a\psi \rightarrow B_a(\varphi \wedge \psi).$$

The converse statement to (C_a) already holds because of monotonicity; hence, when (C_a) holds, individual beliefs fully preserve conjunctiveness.

Because of the powerful monotonicity assumption, our systems are open to the widely discussed objection of 'logical omniscience'.[3] Vardi (1986, 1989) alleviates the problem of logical omniscience by weakening (RM_a) into the familar *rule of equivalence* of so-called classical systems:

$$(\mathrm{RE}_a) \quad \frac{\varphi \leftrightarrow \psi}{B_a\varphi \leftrightarrow B_a\psi}\,.$$

There are a host of technical and conceptual problems with the axiomatization of common belief under (RE_a). These problems explain why we strengthen here (RE_a) into (RM_a). Note that one should not expect all and every property of the B_a to be transmitted to C just by (FP) and (RI). Intuitively, this is one of the reasons why we explicitly require the monotonicity of common belief in (RM_C).

(FP) says, in effect, that common belief implies everybody's belief as well as everybody's belief of the fact that common belief prevails, a property that explains why this axiom is referred to as a *fixed-point* one. (RI) says, in essence, that if a statement is inherently everybody's belief, it is common belief. The label *rule of induction* may be justified as follows. Using (RM_a), $a \in A$ and (Def.E), we note that E is monotonic:

$$(\mathrm{RM}_E) \quad \frac{\varphi \rightarrow \psi}{E\varphi \rightarrow E\psi}\,.$$

A simple inductive argument then leads to

$$\frac{\varphi \rightarrow E\varphi}{E\varphi \rightarrow E^k\varphi}\,, \quad \text{for all } k > 0\,,$$

where E^k is of course $\overbrace{E, \ldots, E}^{k\ times}$. Very intuitively, (RI) says that this inductive process can be recapitulated into a singular inference. (FP) and (RI) are typical examples of the current method of axiomatizing common belief. FHMV have a closely related variant, e.g. Halpern and Moses (1992, Section 4). Whatever the technical differences,[4] the axiomatization of common belief *must* be of the circular kind because of the logician's commitment to finite conjunctions.

Other axioms to be considered in the paper are, for any $a \in A$:

(4_a) $B_a\varphi \rightarrow B_aB_a\varphi$;

(5_a) $\neg B_a\varphi \rightarrow B_a \neg B_a\varphi$;

(T_a) $B_a\varphi \rightarrow \varphi$;

(T'_a) $B_aB_a\varphi \rightarrow B_a\varphi$;

(D_a) $B_a\varphi \rightarrow \neg B_a \neg \varphi$.

[3] See FHMV (forthcoming) for a survey of this problem and its tentative solutions.

[4] See Lismont and Mongin (1994a, Section 3) for a comparison with the FHMV axiom set.

(4_a) and (5_a) are *positive* and *negative introspection* axioms. The latter has come under fierce attack from many writers in epistemic logic and AI.[5] The *truth axiom* (T_a) may be weakened into (T'_a) to deliver a converse statement to the seemingly innocuous (4_a). (D_a) is known to be equivalent to (P_a) when (RM_a), (N_a) and (C_a) hold (see Chellas, 1980, p. 133). Obviously, any modal logic containing (T_a) also contains (D_a).

The axiom set consisting of (RM_a), (N_a) and (C_a) is the K-*system for agent a*, to be denoted by K_a. As is well known, K is the weakest system that can be interpreted by means of Kripke structures. Among the *Normal systems for a* (i.e. the stronger systems than K_a), $K_aD_a = K_aP_a$, $K_aD_a4_a$, and (perhaps not very plausibly) $K_aD_a4_a5_a$ have been defended as formal accounts of individual belief. Similarly, K_aT_a, $K_aT_a4_a$, and (not so obviously) $K_aT_a4_a5_a$, which is equivalent to $K_aT_a5_a$, have been defended as accounts of individual knowledge.

All systems considered in this paper include MC_A. We denote them as $\Sigma_C(\cdot)$. That is to say, $\Sigma_C(K_A)$ is $MC_a + K_a$ for all $a \in A$, and the like. The formal inference relation of system $\Sigma_C(\cdot)$ is denoted by $\vdash_{\Sigma_C(\cdot)}$. The subscript will be omitted when the context is obvious. Our definition of the formal inference relation is the standard one. If $\Gamma \subseteq \Phi$ and $\varphi \in \Phi$, we define $\Gamma \vdash_{\Sigma_C(\cdot)} \varphi$ by the following property: there are a finite number of well-formed formulas $\varphi_1, \ldots, \varphi_n \in \Gamma$ such that $\vdash_{\Sigma_C(\cdot)} \varphi_1 \wedge \cdots \wedge \varphi_n \rightarrow \varphi$ holds.

We end up this section by stating representative theorems that can be derived from MC_A and stronger systems.

Proposition 1.

(i) $\vdash_{MC_A} C\varphi \rightarrow E^k\varphi$, *for all* $k \geq 1$;
(ii) $\vdash_{MC_A} C\varphi \rightarrow C^k\varphi$, *for all* $k \geq 1$;
(iii) $\vdash_{MC_A} C\varphi \leftrightarrow E(\varphi \wedge C\varphi)$;
(iv) $\vdash_{MC_A} C\varphi \rightarrow EC\varphi$;
(v) $\vdash_{MC_A} EC\varphi \rightarrow CE\varphi$;
(vi) $\vdash_{MC_A} C\varphi \rightarrow CE\varphi$;
(vii) $\vdash_{\Sigma_C(C_A)} CE\varphi \rightarrow EC\varphi$;
(viii) $\vdash_{\Sigma_C(C_A)} C\varphi \leftrightarrow EC\varphi \wedge E\varphi$;
(ix) $\vdash_{\Sigma_C(C_A)} C\varphi \leftrightarrow CE\varphi \wedge E\varphi$;
(x) $\vdash_{\Sigma_C(C_A)} C\varphi \leftrightarrow C(E\varphi \wedge \varphi)$;
(xi) $\vdash_{\Sigma_C(C_A)} C\varphi \wedge C\psi \rightarrow C(\varphi \wedge \psi)$;
(xii) $\vdash_{\Sigma_C(N_A)} C\top$;
(xiii) $\vdash_{\Sigma_C(P_A)} \neg C\bot$;
(xiv) $\vdash_{\Sigma_C(T_A)} C\varphi \rightarrow \varphi$;
(xv) $\vdash_{\Sigma_C(T_A)} C\varphi \leftrightarrow C^k\varphi$, *for all* $k \geq 2$;
(xvi) $\vdash_{\Sigma_C(C_A+T'_A)} C\varphi \leftrightarrow C^k\varphi$, *for all* $k \geq 2$.

Sketch of the proof.

(i) $\vdash_{MC_A} C\varphi \rightarrow E\varphi$: from (FP) and ($RM_E$). The theorem for any $k > 1$ follows inductively using (FP) and (RM_E).

(ii) $\vdash_{MC_A} C\varphi \rightarrow CC\varphi$: from (FP), (RI) and ($RM_C$). The theorem for any $k > 2$ follows inductively using (RM_C).

[5] For a recent example, see Modica and Rustichini's (1994) discussion of negative introspection in relation to 'awareness'.

(iii) $\vdash_{\mathrm{MC}_A} E(\varphi \wedge C\varphi) \rightarrow C\varphi$: from (FP), (RI) and ($\mathrm{RM}_C$).
(iv) $\vdash_{\mathrm{MC}_A} C\varphi \rightarrow EC\varphi$: from (FP) and ($\mathrm{RM}_E$).
(v) $\vdash_{\mathrm{MC}_A} EC\varphi \rightarrow CE\varphi$: from (iv), (RI) and ($\mathrm{RM}_C$) as applied to (i) with $k = 1$.
(vi) $\vdash_{\mathrm{MC}_A} C\varphi \rightarrow CE\varphi$: from (iv) and (v).
(vii) $\vdash_{\Sigma_C(\mathrm{C}_A)} CE\varphi \rightarrow EC\varphi$: from

(a)	$CE\varphi \wedge E\varphi \rightarrow ECE\varphi \wedge EE\varphi \wedge E\varphi$	(FP), (RM_E);
(b)	$CE\varphi \wedge E\varphi \wedge \varphi \rightarrow E(CE\varphi \wedge E\varphi \wedge \varphi)$	(a), (C_A);
(c)	$E(CE\varphi \wedge E\varphi \wedge \varphi) \rightarrow C(CE\varphi \wedge E\varphi \wedge \varphi)$	(b), (RI);
(d)	$E(CE\varphi \wedge E\varphi \wedge \varphi) \rightarrow C\varphi$	(c), (RM_C);
(e)	$CE\varphi \wedge E\varphi \rightarrow C\varphi$	(a), (d), (C_A);
(f)	$E(CE\varphi \wedge E\varphi) \rightarrow EC\varphi$	(e), (RM_E);
(g)	$CE\varphi \rightarrow EC\varphi$	(f), (FP).

(viii) $\vdash_{\Sigma_C(\mathrm{C}_A)} C\varphi \leftrightarrow EC\varphi \wedge E\varphi$: from (i), (iv), (iii) and (C_A).
(ix) $\vdash_{\Sigma_C(\mathrm{C}_A)} C\varphi \leftrightarrow CE\varphi \wedge E\varphi$: from (v), (vii) and (viii).
(x) $\vdash_{\Sigma_C(\mathrm{C}_A)} C\varphi \leftrightarrow C(E\varphi \wedge \varphi)$: from (v), ($\mathrm{C}_A$) and ($\mathrm{RM}_C$).
(xi) $\vdash_{\Sigma_C(\mathrm{C}_A)} C\varphi \wedge C\psi \rightarrow C(\varphi \wedge \psi)$: from

(a)	$C\varphi \wedge C\psi \rightarrow E(C\varphi \wedge \varphi \wedge C\psi \wedge \psi)$	(FP), (C_A);
(b)	$C\varphi \wedge C\psi \wedge \varphi \wedge \psi \rightarrow E(C\varphi \wedge C\psi \wedge \varphi \wedge \psi)$	(a), (RM_E);
(c)	$E(C\varphi \wedge C\psi \wedge \varphi \wedge \psi) \rightarrow C(C\varphi \wedge C\psi \wedge \varphi \wedge \psi)$	(b), (RI)
(d)	$C\varphi \wedge C\psi \rightarrow C(\varphi \wedge \psi)$	(c), (RM_C), (a).

(xii) $\vdash_{\Sigma_C(\mathrm{N}_A)} C\top$: from ($\mathrm{N}_A$), (Def.$E$), ($\mathrm{RM}_E$) and (RI).
(xiii) $\vdash_{\Sigma_C(\mathrm{P}_A)} \neg C\bot$: from (i), (Def.$E$) and ($\mathrm{P}_A$).
(xiv) $\vdash_{\Sigma_C(\mathrm{T}_A)} C\varphi \rightarrow \varphi$: from (i), (Def.$E$) and ($\mathrm{T}_A$).
(xv) $\vdash_{\Sigma_C(\mathrm{T}_A)} C\varphi \leftrightarrow C^k\varphi$, for all $k \geq 2$: from (ii), (xiv) and (RM_C).
(xvi) $\vdash_{\Sigma_C(\mathrm{C}_A+\mathrm{T}'_A)} C\varphi \leftrightarrow C^k\varphi$, for all $k \geq 2$: from (ii) and

(a)	$CC\varphi \rightarrow EC\varphi \wedge EE\varphi$	(i), (RM_E);
(b)	$EE\varphi \rightarrow E\varphi$	(Def.E), (RM_E), (T'_A);
(c)	$CC\varphi \rightarrow C\varphi$	(a), (b), (viii);
(d)	$C^k\varphi \rightarrow C\varphi$	by induction on k, using (RM_C). □

This proposition consists of three groups of results. Property (i) formalizes the desirable iterative property of common belief. Properties (ii) to (x) are related to the specifically fixed-point property of common belief. Interestingly, the weakening of $\Sigma_C(\mathrm{K}_A)$ – or, for that matter, $\Sigma_C(\mathrm{C}_A)$ – into MC_A results in the loss of some of the variants. This negative fact can be checked by semantic means once determination theorems are proved. Finally, properties (xi) to (xv) show that the common-belief operator inherits the following properties of individual operators: (C_A), (N_A), (P_A), (T_A), when they are added to MC_A. A particular consequence is that one ipso facto endows C with a K-system by assuming K_A. As (xvi) shows, the transmission of property (T'_A) to C is slightly more complex because it also depends on (C_A).

It should be emphasized that the parallelism between C and the individual operators breaks down completely when it comes to (4_A) and (5_A). As (ii) demonstrates, C obeys positive introspection *even in systems without* (4_A). Using a semantic argument once determination theorems become available, it can be seen that *even in systems involving* (5_A), it is not a theorem that C satisfies negative introspection.

3. A determination theorem for monotonic belief and common belief

We start by introducing our variant of neighborhood semantics, and then prove a determination result in this framework.

The C operator has no explicit counterpart in the structures that we are considering. Hence, these will be of the familiar types in modal logic, barring the minor differences introduced by the multi-agent framework. *A monotonic neighborhood* or *Scott structure* is any $(|A|+2)$-tuple:

$$m = \langle W, (N_a)_{a\in A}, v\rangle ,$$

where

- W is a nonempty set (referred to as a set of possible worlds),
- for any $a \in A$, N_a is a mapping $W \to \mathcal{P}(\mathcal{P}(W))$ – where $\mathcal{P}(\cdot)$ denotes the power set – such that for any $w \in W$, $N_a(w)$ is closed under supersets (i.e. if $P \in N_a(w)$ and $P \subseteq P' \subseteq W$, then $P' \in N_a(w)$);
- v is a mapping $W \times PV \to \{0, 1\}$ (referred to as a valuation).

For convenience, we introduce the mapping $N_E: W \to \mathcal{P}(\mathcal{P}(W))$ defined by

$$N_E(w) = \bigcap_{a\in A} N_a(w) .$$

Except for sentences $\varphi = C\psi$, to be discussed below, our definition of the validation relation for the monotonic neighborhood structures is completely standard. Then, for any $m = \langle W, (N_a)_{a\in A}, v\rangle$:

- if $\varphi \in PV$, $\langle m, w\rangle \models \varphi \Leftrightarrow v(w, \varphi) = 1$;
- if $\varphi = \neg\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow \langle m, w\rangle \not\models \psi$;
- if $\varphi = \psi_1 \wedge \psi_2$, $\langle m, w\rangle \models \varphi \Leftrightarrow \langle m, w\rangle \models \psi_1$ and $\langle m, w\rangle \models \psi_2$,

and similarly for the remaining propositional clauses:

- if $\varphi = B_a\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow [\![\psi]\!]^m \in N_a(w)$;
- if $\varphi = E\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow [\![\psi]\!]^m \in N_E(w)$,

where $[\![\psi]\!]^m$ denotes the truth set of ψ, i.e. $\{w' \in W \mid \langle m, w'\rangle \models \psi\}$ (the superscript m may be omitted).

Henceforth, we shall adhere to the philosophical use of reserving the word 'proposition' for those subsets of possible worlds that are truth sets of some formula.

The neighborhood functions N_a and N_E are easily understood in the present context of epistemic applications of modal logic. They provide systems of belief for a and E ('everybody'), respectively. To every world w they associate the propositions (semantically viewed) that are (semantically) believed at w. (This is not to say that all and every subset in $N_a(w)$, $N_E(w)$ is a proposition.) Mongin (1994) suggests that this semantic rendering of belief is more natural than the more popular one by means of Kripke structures. Part of the attraction of neighborhood structures in epistemic logic stems from the fact that they are a generalization of the Boolean structures of measure theory. Recall also the important point that probabilistic reasoning violates the conjunctiveness axiom (C_a), which is part of Kripke's system.

We proceed now to the unfamiliar component of the semantics. We define a subset P of W to be *belief-closed* (abbreviated as: b.c.) if:

$$\forall w \in P,\ P \in N_E(w)\ .$$

Take a subset P that happens to be a proposition: then the condition of belief closure stipulates that P is believed in every world where it is true.

The validation clause for common belief may now be introduced:

- if $\varphi = C\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow \exists P \in N_E(w)$ such that $P \subseteq [\![\psi]\!]^m$ and P is b.c.

Again, the intuitive motivation of this clause becomes clear when one restricts one's consideration to propositions. Then, the definition simply states that common belief of ψ prevails at w if the proposition corresponding to ψ includes a (possibly different) proposition that is both believed by everybody at w and belief-closed.

Let us denote the class of monotonic neighborhood structures by $\mathcal{M}^N$. As usual, $m \models \varphi$ and $\mathcal{M} \models \varphi$ abbreviate [for all $w \in W$, $\langle m, w\rangle \models \varphi$] and [for all m in the relevant class $\mathcal{M}$, $m \models \varphi$], respectively. Reference to $\mathcal{M}$ may be omitted when the context is obvious. As usual, again, given $\Sigma \subseteq \Phi$, $\Sigma \models \varphi$ means the following: for all structures m in the underlying class $\mathcal{M}$, if [for all $\psi \in \Sigma$, $m \models \psi$], then $m \models \varphi$. The definitions of this paragraph will also be used in the context of Kripke structures.

We may now state the result of this section:

Theorem 2.

$$\vdash_{\mathrm{MC}_A} \varphi \Leftrightarrow \mathcal{M}^N \models \varphi\ .$$

The proofs of this and the following theorems will be simplified by introducing further notation. Given $m \in \mathcal{M}^N$, consider the mapping b_a: $\mathcal{P}(W) \to \mathcal{P}(W)$:

$$b_a(P) = \{w \in W \mid P \in N_a(w)\}\ .$$

It satisfies the condition that, for all formula φ,

$$[\![B_a\varphi]\!]^m = b_a([\![\varphi]\!]^m)\ .$$

Similarly, take the mapping $b_E: \mathcal{P}(W) \rightarrow \mathcal{P}(W)$, which is defined by either:

$$b_E(P) = \{w \in W \mid P \in N_E(w)\} ,$$

or (equivalently) by

$$b_E(P) = \bigcap_{a \in A} b_a(P) .$$

It satisfies the property that, for any formula φ,

$$[\![E\varphi]\!]^m = b_E([\![\varphi]\!]^m) .$$

Note carefully that b_a, b_E are monotonic from the construction of $\mathcal{M}^N$.[6]

Now, the following easy lemma holds for any $m \in \mathcal{M}^N$:

Lemma 1. (1) *$P \in \mathcal{P}(W)$ is belief-closed if and only if $P \subseteq b_E(P)$.*

(2) *$\langle m, w \rangle \models C\varphi \Leftrightarrow \exists P \subseteq [\![\varphi]\!]^m \cap b_E(P)$ such that $w \in b_E(P)$.*

(3) $[\![C\varphi]\!]^m = \bigcup \{b_E(P) \mid P \subseteq [\![\varphi]\!]^m \cap b_E(P)\}$.

Proof of the soundness part of Theorem 2. Obviously, (RM_A), (Def.E) and (RM_C) are valid in $\mathcal{M}^N$. We have to check that (FP) as well is valid, i.e. that for any model:

$$[\![C\varphi]\!] \subseteq [\![E(C\varphi \wedge \varphi)]\!] .$$

Using the restatement of $[\![C\varphi]\!]$ in Lemma 1, take any $P \subseteq W$ such that $P \subseteq [\![\varphi]\!] \cap b_E(P)$. Then,

$$b_E(P) \subseteq [\![C\varphi]\!] ,$$

and from the monotonicity property of b_E applied twice over:

$$b_E(P) \subseteq b_E([\![\varphi]\!] \cap b_E(P)) \subseteq b_E([\![\varphi]\!] \cap [\![C\varphi]\!]) = [\![E(C\varphi \wedge \varphi)]\!] .$$

Hence the desired inclusion.

In order to check that (RI) is valid we assume that in any model:

$$[\![\varphi]\!] \subseteq [\![E\varphi]\!] ,$$

and aim at showing that

$$[\![E\varphi]\!] \subseteq [\![C\varphi]\!] .$$

From the assumption, $[\![\varphi]\!] = [\![\varphi]\!] \cap b_E([\![\varphi]\!])$. Hence, from the expression of $[\![C\varphi]\!]$ in Lemma 1, $b_E([\![\varphi]\!]) \subseteq [\![C\varphi]\!]$ holds, which is the desired conclusion. □

[6] Contrary to the *syntactical* operators B_a, E, these set-theoretic operators are well known to economists and game-theorists; see, for example, Bacharach (1985) and Geanakoplos (1992). Notice that they are normally endowed with stronger properties than just monotonicity.

Proof of the completeness part of Theorem 2. It involves four lemmas. The general idea is to prove the relevant implication:

$$\mathcal{M}^N \models \psi \Rightarrow \vdash_{\mathrm{MC}_A} \psi \,,$$

for a *given* formula ψ, adapting the standard construction of a canonical model based on the maximal consistent sets of formulas.[7] Definitions and facts relative to the maximal consistent subsets of formulas, as in Chellas (1980, Chapter 2) are taken for granted in this paper.

Given any $\varphi \in \Phi$, we inductively define the *depth* of φ, denoted by $\mathrm{dp}(\varphi)$:

- if $\varphi \in PV$, then $\mathrm{dp}(\varphi) = 0$;
- $\mathrm{dp}(\neg\varphi) = \mathrm{dp}(\varphi)$;
- $\mathrm{dp}(\varphi \vee \psi) = \mathrm{dp}(\varphi \wedge \psi) = \mathrm{dp}(\varphi \rightarrow \psi) = \mathrm{dp}(\varphi \leftrightarrow \psi) = \max(\mathrm{dp}(\varphi), \mathrm{dp}(\psi))$;
- $\mathrm{dp}(B_a\varphi) = \mathrm{dp}(E\varphi) = \mathrm{dp}(C\varphi) + 1$.

We also define $\Phi[\psi]$ to be the subset of those formulas that satisfy the following two conditions:

(i) they are constructed from the subset of propositional variables occurring in ψ;

(ii) they have depth at most $\mathrm{dp}(\psi)$.

A set $\Gamma \subseteq \Phi[\psi]$ is said to be *$\Phi[\psi]$-maximal consistent* if it is consistent and no formula of $\Phi[\psi]$ can be added to it without making it inconsistent. Clearly, the set I^ψ of the $\Phi[\psi]$-maximal consistent sets is exactly the set of intersections $\Gamma' \cap \Phi[\psi]$, whenever Γ' ranges over the set of maximal consistent subsets of Φ. In the definition below of a $\Phi[\psi]$-canonical model, we shall take the set of possible worlds W to be I^ψ.

For any $\varphi \in \Phi[\psi]$, the notation $[\varphi]^\psi$ will refer to the set $\{\Gamma \in I^\psi \mid \varphi \in \Gamma\}$, or, equivalently, $\{\Gamma \in I^\psi \mid \Gamma \vdash_{\mathrm{MC}_A} \varphi\}$. Normally, superscripts will be dropped. We leave it for the reader to check the following properties:

Lemma 2. (1) *The set I^ψ is finite.*

(2) *All subsets of I^ψ can be characterized by a formula of $\Phi[\psi]$ in the following sense:*

For any $P \subseteq I^\psi$, there is $\varphi_0 \in \Phi[\psi]$ such that $P = [\varphi_0]$.

(3) *For all $\varphi_1, \varphi_2 \in \Phi[\psi]$, $[\varphi_1] \subseteq [\varphi_2] \Leftrightarrow \vdash \varphi_1 \rightarrow \varphi_2$.*

[7] Completeness proofs of the 'each formula' kind have already been devised in modal logic to solve altogether different problems from those raised by the C operator: see Boolos (1979) or Cresswell (1983). Another source is the selective filtration method used in dynamic logic, e.g. Harel (1984) or Goldblatt (1987). All existing completeness results relative to common belief depend on the 'each formula' type of proof: see Kraus and Lehmann (1988), Halpern and Moses (1990, 1992), Lismont (1993), and Lismont and Mongin (1994b).

The neighborhood functions of the $\Phi[\psi]$-canonical model are constructed as:

$$\forall \Gamma \in I^{\psi},\ N_a(\Gamma) \stackrel{\text{def}}{=} \{P \subseteq I^{\psi} \mid \exists \varphi \in \Phi[\psi] \text{ such that } [\varphi] \subseteq P \text{ and } \Gamma \vdash B_a\varphi\}\ .$$

Hence the associated functions $b_a : \mathcal{P}(I^{\psi}) \to \mathcal{P}(I^{\psi})$ are

$$b_a(P) \stackrel{\text{def}}{=} \{\Gamma \in I^{\psi} \mid P \in N_a(\Gamma)\} = \{\Gamma \in I^{\psi} \mid \exists \varphi \in \Phi[\psi] \text{ such that } [\varphi] \subseteq P \text{ and } \Gamma \vdash B_a\varphi\}\ .$$

Lemma 3. (1) *For all* $\varphi \in \Phi[\psi]$, $b_a([\varphi]) = \{\Gamma \in I^{\psi} \mid \Gamma \vdash B_a\varphi\}$ *and* $b_E([\varphi]) = \{\Gamma \in I^{\psi} \mid \Gamma \vdash E\varphi\}$.

(2) *For all* $\varphi \in \Phi[\psi]$ *such that* $\text{dp}(\varphi) < \text{dp}(\psi)$, $b_a([\varphi]) = [B_a\varphi]$ *and* $b_E([\varphi]) = [E\varphi]$.

Proof. For the first assertion, the inclusion $\{\Gamma \in I^{\psi} \mid \Gamma \vdash B_a\varphi\} \subseteq b_a([\varphi])$ is trivial. To check the opposite inclusion, take any Γ such that $\exists \varphi' \in \Phi[\psi]$ satisfying $[\varphi'] \subseteq [\varphi]$ and $\Gamma \vdash B_a\varphi'$. The first condition and Lemma 2 yield that $\vdash \varphi' \to \varphi$. Using the second condition and (RM_a), we conclude that $\Gamma \vdash B_a\varphi$. Hence, $b_a([\varphi]) \subseteq \{\Gamma \in I^{\psi} \mid \Gamma \vdash B_a\varphi\}$. The second equation in the first statement immediately follows from the first and axiom (Def.E).

The second assertion follows from the first and the fact that $\Gamma \vdash \varphi'$ and $\varphi' \in \Gamma$ are equivalent when $\varphi' \in \Phi[\psi]$. □

Given $\psi \in \Phi$ and $\Phi[\psi]$, a *$\Phi[\psi]$-canonical model* is a *$(|A| + 2)$-tuple*:

$$m^{\psi} = \langle I^{\psi}, (N_a)_{a \in A}, v \rangle\ ,$$

where all symbols but the last one have already been defined. The valuation v is the following function, for all $\Gamma \in I^{\psi}$, for all $p \in PV$:

$$v(\Gamma, p) = 1 \Leftrightarrow p \in \Gamma\ .$$

The $\Phi[\psi]$-canonical models are well-defined members of $\mathcal{M}^N$. In our ψ-dependent approach, Lemma 5 below states a property analogous to that of the canonical models in classical modal logic. Before deriving this lemma, we state another purely syntactical fact:

Lemma 4. *Let* $\varphi \in \Phi[\psi]$ *such that* $\text{dp}(\varphi) < \text{dp}(\psi)$.

(1) *Let* $\Phi \in I^{\psi}$. *Then*

$$C\varphi \in \Gamma \Leftrightarrow \exists \varphi_0 \in \Phi[\psi], \textit{ such that } \Gamma \vdash E\varphi_0,\ \vdash \varphi_0 \to \varphi \quad \textit{and} \quad \vdash \varphi_0 \to E\varphi_0\ .$$

(2) $[C\varphi] = \bigcup \{b_E([\varphi_0]) \mid \varphi_0 \in \Phi[\psi], [\varphi_0] \subseteq [\varphi] \cap b_E([\varphi_0])\}$.

Proof. The second assertion will immediately follow from the first one, Lemma 2 and Lemma 3. For the first assertion, assume that $\varphi \in \Phi[\psi]$ and $\text{dp}(\varphi) < \text{dp}(\psi)$.

Then clearly $\varphi_0 = C\varphi \wedge \varphi \in \Phi[\psi]$. The condition that $\Gamma \vdash E\varphi_0$ results from the assumption that $C\varphi \in \Gamma$ along with (FP). The condition that $\vdash \varphi_0 \rightarrow \varphi$ is trivially met. Finally, $\vdash \varphi_0 \rightarrow E\varphi_0$ also follows from (FP).

Conversely, assume that there is φ_0 as stipulated. (RI) and (RM_C) lead to $\vdash E\varphi_0 \rightarrow C\varphi_0$ and $\vdash C\varphi_0 \rightarrow C\varphi$, respectively. Hence, $\Gamma \vdash C\varphi$, i.e. $C\varphi \in \Gamma$ from the assumption on the depth of φ. □

Lemma 5. *For all* $\varphi \in \Phi[\psi]$, $[\![\varphi]\!]^{m^\psi} = [\varphi]^\psi$.

Proof. The proof goes by induction on the complexity of φ.

- If $\varphi \in PV$, $[\![\varphi]\!] = [\varphi]$ from the definition of v.
- If $\varphi = \neg\varphi'$:

$$
\begin{aligned}
[\![\neg\varphi']\!] &= I^\psi \setminus [\![\varphi']\!] && \text{(from the definition of } \models) \\
&= I^\psi \setminus [\varphi'] && \text{(from the induction hypothesis)} \\
&= [\neg\varphi'] && \text{(from the maximality property)}\,.
\end{aligned}
$$

- The remaining propositional cases are dealt with similarly.
- Suppose that $\varphi = B_a\varphi'$.

$$
\begin{aligned}
[\![B_a\varphi']\!] &= b_a([\![\varphi']\!]) && \text{(from the definitions of } \models \text{ and } b_a) \\
&= b_a([\varphi']) && \text{(from the induction hypothesis)} \\
&= [B_a\varphi'] && \text{(from Lemma 3(2))}\,.
\end{aligned}
$$

- A parallel argument applies to the case in which $\varphi = E\varphi'$.
- Suppose that $\varphi = C\varphi'$.

$$
\begin{aligned}
[\![C\varphi']\!] &= \bigcup \{b_E(P) \mid P \subseteq [\![\varphi']\!] \cap b_E(P)\} && \text{(from Lemma 1)} \\
&= \bigcup \{b_E(P) \mid P \subseteq [\varphi'] \cap b_E(P)\} && \text{(from the induction hypothesis)} \\
&= \bigcup \{b_E([\varphi_0]) \mid \varphi_0 \in \Phi[\psi] \text{ and } [\varphi_0] \subseteq [\varphi'] \cap b_E([\varphi_0])\} && \\
& && \text{(from Lemma 2)} \\
&= [C\varphi'] && \text{(from Lemma 4(2)).} \quad \square
\end{aligned}
$$

End of the proof of Theorem 2. For any given $\psi \in \Phi$, assume that $\mathcal{M}^N \models \psi$. In particular, ψ is valid in m^ψ. Lemma 5 implies that $[\psi]^\psi = [\![\psi]\!]^{m^\psi}$, i.e. $[\psi]^\psi = I^\psi$; hence, that $\psi \in \Gamma$ for all $\Phi[\psi]$-maximal consistent sets Γ. This is equivalent to saying that $\psi \in \Gamma' \cap \Phi[\psi]$ for all maximal consistent sets Γ', i.e. that ψ is in all maximal consistent sets of system MC_A. Hence $\vdash \psi$. □

4. Determination theorems for normal systems of belief and common belief

We now proceed to the Kripkean variant of our semantics. It will be seen to deliver simple determination theorems for $\Sigma_C(\mathrm{K}_A)$ (Theorem 3), as well as for some Normal systems of belief and common belief (Theorem 4).

In the multi-agent framework of this paper, a *Kripke structure* is any $(|A|+2)$-tuple:

$$m = \langle W, (R_a)_{a\in A}, v\rangle ,$$

where W and v are as before, and for any $a \in A$, R_a is a binary relation on W. For convenience we introduce the binary relation R_E defined as $\bigcup_{a\in A} R_a$. The class of Kripke structures will be denoted by $\mathcal{M}^K$.

Except for sentences $\varphi = C\psi$, our definition of the validation relation in $\mathcal{M}^K$ is the familiar one. That is to say, for any $m = \langle W, (R_a)_{a\in A}, v\rangle$, the propositional clauses are as in Section 3, and

- if $\varphi = B_a\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow R_a[w] \subseteq [\![\varphi]\!]^m$,
- *if* $\varphi = E\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow R_E[w] \subseteq [\![\varphi]\!]^m$,

where $R_a[w] \overset{\text{def}}{=} \{w' \in W \mid wR_aw'\}$ and $R_E[w]$ is similarly defined.

In an epistemic context the *accessibility relations* R_a, R_E may be viewed as connecting with $w \in W$ those $w' \in W$ that a, E, respectively, regard as accessible from w. This informal interpretation is rather shaky. As an explicans of belief, the notion of subjective accessibility is obscure. It seems to result from a simple, ad hoc modification in the modal understanding of the relations (as describing objective or metaphysical possibility). We have already argued that it seems more natural to use neighborhood structures than Kripke structures in a context of epistemic applications, especially if one is concerned with connecting one's modelling of belief with the probabilistic one. There are two reasons, however, why it is useful to extend the investigation of this paper to Kripke structures. For one, they have been used extensively, whether explicitly or not, in the semantic discussions of common belief. For another, they provide elegant counterparts to the axioms when MC_A is enriched with further constraints on individual belief, as in $\Sigma_C(\mathrm{K}_A4_A)$ or $\Sigma_C(\mathrm{K}_A\mathrm{T}_A5_A)$.

Given $m \in \mathcal{M}^K$, a set $P \subseteq W$ is said to be *belief-closed* (b.c.) if:

$$\forall w \in P,\ R_E[w] \subseteq P .$$

The validation clause for common belief may now be stated:

- if $\varphi = C\psi$, $\langle m, w\rangle \models \varphi \Leftrightarrow \exists P \subseteq W$ such that $P \subseteq [\![\psi]\!]^m$, $R_E[w] \subseteq P$ and P is b.c.

Following a well-known construction of modal logic, $\mathcal{M}^K$ can be embedded into $\mathcal{M}^N$ (see Chellas, 1980, Chapters 7 and 8). Using this construction, it is routine to check that the above definitions of belief closure and the validation clause for $C\psi$ are equivalent to those already given in Section 3, when $\mathcal{M}^N$ is restricted to $\mathcal{M}^K$.

Given $m \in \mathcal{M}^K$, we shall introduce the mappings $b_a, b_E : \mathcal{P}(W) \to \mathcal{P}(W)$ defined by

$$b_a(P) = \{w \in W \mid R_a[w] \subseteq P\} ,$$
$$b_E(P) = \{w \in W \mid R_E[w] \subseteq P\} ,$$

or equivalently,

$$b_E(P) = \bigcap_{a \in A} b_a(P) .$$

The mappings satisfy the property that, for all formulas φ,

$$[\![B_a\varphi]\!]^m = b_a([\![\varphi]\!]^m) \quad \text{and} \quad [\![E\varphi]\!]^m = b_E([\![\varphi]\!]^m) .$$

Clearly, b_a and b_E are monotonic functions. More than that is true: we leave it for the reader to check that Lemma 1 holds again under the present definitions.

Here is a first determination result:

Theorem 3.

$$\vdash_{\Sigma_C(\mathrm{K}_A)} \varphi \Leftrightarrow \mathcal{M}^K \models \varphi .$$

Proof of the soundness part. (Def.E) is obviously valid. The Kripkean features of the semantics mean that (RM_A), (N_A) and (C_A) are valid. In view of the fact that Lemma 1 holds and b_E is monotonic, the proof in Section 3 that (FP) and (RI) are valid is unchanged. □

As it turns out, the use of the b_a and b_E also greatly simplifies the proof of the completeness part. The desired implication,

$$\mathcal{M}^K \models \psi \Rightarrow \vdash_{\Sigma_C(\mathrm{K}_A)} \psi ,$$

will be proved for the *given* formula ψ. Assuming now that $\vdash$ is $\vdash_{\Sigma_C(\mathrm{K}_A)}$, we may rely on the earlier definitions of $\Phi[\psi]$, I^ψ and $[\varphi]^\psi$.

We leave it for the reader to check that Lemma 2 holds of system $\Sigma_C(\mathrm{K}_A)$.

A $\Phi[\psi]$-*canonical model* is now a $(|A| + 2)$-*tuple*:

$$m^\psi = \langle I^\psi, (R_a)_{a \in A}, v \rangle ,$$

where, for all $\Gamma \in I^\psi$,

- $R_a[\Gamma] = \{\Delta \in I^\psi \mid \{\varphi \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi\} \subseteq \Delta\}$;
- for all $\varphi \in PV$, $v(\Gamma, \varphi) = 1 \Leftrightarrow \varphi \in \Gamma$.

The mappings b_a relative to the $\Phi[\psi]$-canonical model are: for all $P \in \mathcal{P}(I^\psi)$, for all $\Gamma \in I^\psi$,

$$\Gamma \in b_a(P) \Leftrightarrow \forall \Delta \in I^\psi , \quad \text{if } \{\varphi \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi\} \subseteq \Delta , \quad \text{then } \Delta \in P .$$

We now check that Lemma 3 of Section 3 holds true of these newly defined b_a:

Lemma 6. (1) *For all* $\varphi \in \Phi[\psi]$, $b_a([\varphi]) = \{\Gamma \in I^\psi \mid \Gamma \vdash B_a\varphi\}$ *and* $b_E([\varphi]) = \{\Gamma \in I^\psi \mid \Gamma \vdash E\varphi\}$.

(2) *For all* $\varphi \in \Phi[\psi]$, *such that* $\mathrm{dp}(\varphi) < \mathrm{dp}(\psi)$, $b_a([\varphi]) = [B_a\varphi]$ *and* $b_E([\varphi]) = [E\varphi]$.

Proof. In order to prove the first equation in the first assertion, we have to check that the following two statements are equivalent:

(a) for all $\Delta \in I^\psi$, if $\{\varphi' \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi'\} \subseteq \Delta$ then $\varphi \in \Delta$,

(b) $\Gamma \vdash B_a\varphi$.

It is trivial that (b) implies (a). For the converse, assume that (a) is true. The formula φ is in all $\Phi[\psi]$-maximal consistent sets containing $\{\varphi' \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi'\}$. This is equivalent to saying that

$$\{\varphi' \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi'\} \vdash \varphi \,.$$

Hence, there exist formulas $\varphi_1, \ldots, \varphi_n$ $(n \geq 0)$ such that

(i) for all $i \leq n$, $\Gamma \vdash B_a\varphi_i$;

(ii) $\vdash \varphi_1 \wedge \cdots \wedge \varphi_n \to \varphi$.

If $n = 0$, (ii) is simply $\vdash \varphi$, and (N_a) and (RM_a) imply that $\Gamma \vdash B_a\varphi$. If $n > 0$, the application of (RM_a) and (C_a) leads to the same conclusion.

The end of the proof is exactly the same as in Lemma 3. □

End of the Proof of Theorem 3. The proofs of Lemma 4, Lemma 5 and the completeness part of Theorem 2 of Section 3 can be repeated word for word. □

We now proceed to derive determination results for some Normal systems for A. The schemata (T_a), (D_a), (4_a) and (5_a) will be seen to have their usual relational counterparts in the Kripke semantics (see Chellas, 1980, Chapter 5, for definitions). In our multi-agent framework a reflexive (serial, transitive, reflexive-Euclidean) Kripke structure is one in which, for every a, R_a is reflexive (serial, transitive, reflexive and Euclidean, respectively). (Recall that a reflexive and Euclidean relation is an equivalence relation, and conversely.)

Lemma 7. (1) *Let* R^* *be the transitive closure of a binary relation* R *on* W *and* $P \subseteq W$:

$$(\forall w \in P, R[w] \subseteq P) \Leftrightarrow (\forall w \in P, R^*[w] \subseteq P) \,.$$

(2) *For any* Γ *and* $\Delta \in I^\psi$ *such that* $\{\varphi \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi\} \subseteq \Delta$, *there exists a maximal consistent set* $\Gamma' \supseteq \Gamma$ *such that* $\{\varphi \in \Phi \mid B_a\varphi \in \Gamma'\} \cup \Delta$ *is consistent.*

Proof. The first point is obvious. For the second, notice first that there are a finite number of classes in Δ relative to logical equivalence. In any class, select a formula and call φ_* the chosen formula in the class of φ. Suppose now that for every maximal consistent set $\Gamma' \supseteq \Gamma$, $\{\varphi \in \Phi \mid B_a\varphi \in \Gamma'\} \cup \Delta$ is not consistent. Then, for each Γ', there exists $\varphi^{\Gamma'}$ with $B_a\varphi^{\Gamma'} \in \Gamma'$, and $\psi^{\Gamma'} \in \Delta$ such that $\vdash \varphi^{\Gamma'} \to \neg\psi_*^{\Gamma'}$. Defining ψ_* to be the (finite) conjunction of the formulas $\psi_*^{\Gamma'}$, $\vdash \varphi^{\Gamma'} \to \neg\psi_*$ holds, and applying (RM_a), $\vdash B_a\varphi^{\Gamma'} \to B_a \neg \psi_*$. This implies that for all maximal consistent sets $\Gamma' \supseteq \Gamma$, $B_a \neg \psi_* \in \Gamma'$, i.e. $\Gamma \vdash B_a \neg \psi_*$. From the assumption and the fact that $\psi_* \in \Phi[\psi]$, $\neg\varphi_* \in \Delta$, a contradiction. □

Theorem 4. *$\Sigma_C(K_A T_A)$, $\Sigma_C(K_A D_A)$, $\Sigma_C(K_A 4_A)$ and $\Sigma_C(K_A T_A 5_A)$ are determined by the classes of reflexive, serial, transitive and reflexive-Euclidean Kripke structures, respectively.*

Proof. Concerning (T_A). The only property to be checked is that when (T_a) holds for $a \in A$, the $\Phi[\psi]$-canonical model $m^\psi = \langle I^\psi, (R_a)_{a\in A}, v\rangle$ is reflexive. Then the previous proof can be repeated without change. For any $\Gamma \in I^\psi$, consider the set:

$$\{\varphi \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi\}\ .$$

From (T_a) this set is included in Γ. Hence, $\Gamma \in R_a[\Gamma]$, the desired property.

Concerning (D_A). Similarly, we check that these axioms, or, equivalently, (P_A), make the $\Phi[\psi]$-canonical model a serial one. We have to prove that for any $a \in A$ and any $\Gamma \in I^\psi$, there is $\Delta \in I^\psi$ such that

$$\{\varphi \in \Phi[\psi] \mid \Gamma \vdash B_a\varphi\} \subseteq \Delta\ .$$

Because of (N_a), the set on the left is never empty. Lindenbaum's lemma implies that there is Δ as required unless this set proves a contradiction. Assume that it does. Then, there would be $\varphi_1, \ldots, \varphi_n$ such that (i) $\vdash \varphi_1 \wedge \cdots \wedge \varphi_n \to \perp$ and (ii) $\Gamma \vdash B_a\varphi_1 \wedge \cdots \wedge B_a\varphi_n$. Applying (RM_a) and (C_a) to (i), we derive $\vdash B_a(\varphi_1 \wedge \cdots \wedge \varphi_n) \to B_a\perp$. Then, from (ii), $\Gamma \vdash B_a\perp$, which contradicts (P_a).

Concerning (4_A). Define new canonical models as follows.[8]

$$m_*^\psi = \langle I^\psi, (R_a^*)_{a\in A}, v\rangle\ ,$$

where I^ψ, R_a and v are as above, and R_a^* is the transitive closure of R_a.

[8] Those readers who know the standard techniques of modal logic might be puzzled by the roundabout derivation below of completeness for $\Sigma_C(\mathrm{K}_A 4_A)$ and $\Sigma_C(\mathrm{K}_A \mathrm{T}_A 5_A)$. It would seem as if we could prove the completeness of *any* system stronger than $\Sigma_C(\mathrm{K}_A)$ just by checking the relational properties of the corresponding $\Phi[\psi]$-canonical models. This simple intuition works in the $\Sigma_C(\mathrm{K}_A \mathrm{T}_A)$ and $\Sigma_C(\mathrm{K}_A \mathrm{D}_A)$ cases. But it becomes difficult to apply to systems involving axiom schemata (4_A) and (5_A). The intuitive reason is that one has now to consider formulas whose depth exceeds that of the given ψ, and accordingly modify the initial notion of a $\Phi[\psi]$-canonical model. This is why we adopted a nonstandard proof technique to deal with the $\Sigma_C(\mathrm{K}_A 4_A)$ and $\Sigma_C(\mathrm{K}_A \mathrm{T}_A 5_A)$ cases.

Obviously, these models are transitive. In order to prove completeness, we have to check by induction that for every $a \in A$ and all $\varphi \in \Phi[\psi]$:

$$[\![\varphi]\!]^{m_*^\psi} = [\varphi] .$$

The proof of Lemma 5 should be adapted only for the case in which $\varphi = B_a\varphi'$ and $\varphi = C\varphi'$, where $\varphi' \in \Phi[\psi]$ and $\mathrm{dp}(\varphi') < \mathrm{dp}(\psi)$.

In the case in which $\varphi = B_a\varphi'$, the proof that $b_a^*([\varphi']) = b_a([\varphi'])$, or, equivalently, that for all $\Gamma \in I^\psi$:

$$R_a[\Gamma] \subseteq [\varphi'] \Leftrightarrow R_a^*[\Gamma] \subseteq [\varphi'] ,$$

will imply the conclusion. The inclusion from right to left is obvious. To prove the converse inclusion, let us show that

$$R_a[\Gamma] \subseteq [\varphi'] \Rightarrow R_a^n[\Gamma] \subseteq [\varphi'] ,$$

where

$$\Delta \in R_a^n[\Gamma] \Leftrightarrow \exists \Gamma_0, \ldots, \Gamma_n, \text{ such that } \Gamma_{i+1} \in R_a[\Gamma_i],\ \Gamma_0 = \Gamma \text{ and } \Gamma_n = \Delta .$$

Assume that $\Delta \in R_a^n[\Gamma]$ and $R_a[\Gamma] \subseteq [\varphi']$, which is equivalent to $\Gamma \in b_a([\varphi'])$. By Lemma 6, $b_a([\varphi']) = [B_a\varphi']$. Thus, $B_a\varphi' \in \Gamma$. By (4_a), $\Gamma \vdash B_aB_a\varphi'$, which implies that $B_a\varphi' \in \Gamma_1$. Applying (4_a) repeatedly, we conclude that $B_a\varphi' \in \Gamma_{n-1}$ and $\varphi' \in \Gamma_n = \Delta$.

In the case in which $\varphi = C\varphi'$, the proof that [for all $P \subseteq I^\psi$, $P \subseteq b_E(P)$] is equivalent to $[P \subseteq b_E^*(P)]$ will deliver the conclusion. This fact follows from Lemma 7(1). We may apply it since the transitive closure of the union of the transitive closures R_a^* of the individual relations R_a is equal to the transitive closure of the union of the individual relations R_a.

Concerning (T_A5_A). Define the canonical models m_*^ψ as in the previous case. They are reflexive because of (T_A) and transitive from the construction. As (4_A) is derivable from (T_A) and (5_A), the proof of the last paragraph applies. It remains to show that canonical models are symmetric. We have to prove that for all $a \in A$ and all $\Gamma, \Delta \in I^\psi$:

$$(\forall \varphi \in \Phi[\psi],\ \Gamma \vdash B_a\varphi \Rightarrow \varphi \in \Delta) \Rightarrow (\forall \varphi \in \Phi[\psi],\ \Delta \vdash B_a\varphi \Rightarrow \varphi \in \Gamma) .$$

Assume the antecedent and take φ such that $\varphi \in \Phi[\psi]$ and $\varphi \not\in \Gamma$. Let Γ' be any maximal consistent set such that $\Gamma' \supseteq \Gamma$. Since $\neg\varphi \in \Gamma$, $\varphi \not\in \Gamma'$ and from (T_a) $B_a\varphi \not\in \Gamma'$, i.e. $\neg B_a\varphi \in \Gamma'$. Applying (5_a) leads to $B_a\neg B_a\varphi \in \Gamma'$. From the assumption and Lemma 7(2), there exists a maximal consistent set $\Gamma' \supseteq \Gamma$ such that $\{\varphi \in \Phi \mid B_a\varphi \in \Gamma'\} \cup \Delta$ is consistent. Lindenbaum's lemma then ensures that there exists a maximal consistent set $\Delta' \supseteq \{\varphi \in \Phi \mid B_a\varphi \in \Gamma'\} \cup \Delta$. As $B_a\neg B_a\varphi \in \Gamma'$, $\neg B_a\varphi \in \Delta'$ and $\Delta \nvdash B_a\varphi$. $\square$

5. Fixed-point versus iterate validation clauses, and some further comments

The primary aim of this section is to analyze the relation of the belief closure semantics to alternative semantics, i.e. those mentioned under headings (1) and (3) in the introduction. We first note that the belief closure semantics satisfies the following basic property, to be referred to as the Minimum Semantic Requirement: if common belief of φ holds in a world, then shared belief of φ, shared belief of shared belief of φ, etc. hold in that world. This property is, of course, the semantic counterpart of (i) in Proposition 1. In the case of Kripke structures, the one-way implication stated in the Minimum Semantic Requirement turns out to be an equivalence. As a consequence, validation clause (1) and the b.c. validation clause (2) are themselves equivalent.

Proposition 5. *For any $m \in \mathcal{M}^N$ and $w \in W$, if $\langle m, w\rangle \models C\varphi$, then $\langle m, w\rangle \models E^k\varphi$, for all $k \geqslant 1$. For any $m \in \mathcal{M}^K$ and $w \in W$, a corresponding implication holds, and, conversely, if $\langle m, w\rangle \models E^k\varphi$, for all $k \geqslant 1$, then $\langle m, w\rangle \models C\varphi$.*

Proof. The proof of the first implication goes by induction on $k \geqslant 1$. Take $m \in \mathcal{M}^N$ such that $\langle m, w\rangle \models C\varphi$. Then there is P as stipulated in the definition. For $k = 1$, using the monotonic closure of $N_E(w)$, we conclude that $[\![\varphi]\!] \in N_E(w)$, i.e. $\langle m, w\rangle \models E\varphi$. Assume that for any $w' \in W$, if $\langle m, w'\rangle \models C\varphi$, then $\langle m, w'\rangle \models E^k\varphi$. It is immediate that for any $w'' \in P$, $\langle m, w''\rangle \models C\varphi$. So $P \subseteq [\![C\varphi]\!] \subseteq [\![E^k\varphi]\!]$. We conclude that $\langle m, w\rangle \models E^{k+1}\varphi$ from the monotonic closure of N_E.

Now, take any $m \in \mathcal{M}^K$ and w in it. The implication just proved holds again since $\mathcal{M}^K$ can be embedded into $\mathcal{M}^N$. For the converse implication, assume that $\langle m, w\rangle \models E^k\varphi$, for all $k \geqslant 1$, and put $P = \bigcap_{k \geqslant 1} [\![E^k\varphi]\!] \cap [\![\varphi]\!]$. Trivially, $P \subseteq [\![\varphi]\!]$. It is easy to check that $R_E[w] \subseteq P$ and P is b.c., so that $\langle m, w\rangle \models C\varphi$. $\square$

We now discuss the type (3) approach to the validation clause for $C\varphi$. If the semantics is Kripkean, it is enough to construct the relational counterpart R_C of C as the *transitive closure* of the shared belief relation R_E,[9] and then state the validation clause for $C\psi$ in usual Kripkean terms:

$$\langle m, w\rangle \models C\psi \overset{\text{def}}{\Leftrightarrow} [\![\psi]\!]^m \subseteq R_C[w]\,.$$

Using this notion of the validation clause, the following observation is trivial:

$$\langle m, w\rangle \models C\psi \Leftrightarrow \forall k \geqslant 1\,, \quad \langle m, w\rangle \models E^k\psi\,.$$

Hence, the type (3) approach collapses into (1), which has just been proved to be equivalent to the b.c. approach.

[9] Formally, if $m = \langle W, (R_a)_{a \in A}, v\rangle \in \mathcal{M}^K$ and $w, w' \in W$, wR_Cw' holds iff there is $n > 1$ and a sequence of n worlds $w_1, \ldots, w_n \in W$ such that $w_1 = w$, $w_n = w'$ and $w_iR_Ew_{i+1}$, for $1 \leqslant i \leqslant n-1$.

The analysis of (3) is not so obvious in the neighborhood context as it is in Kripke's. The validation clause for $C\psi$ can only be defined after the following construction – due to Lismont (1993) – has been made. The general idea is to define N_C as the greatest fixed point of a set-theoretic operator that is defined in terms of the N_a (or N_E) and iterated infinitely. The iteration of the set-theoretic operator is formalized as a transfinite induction. We shall here briefly restate this construction in equivalent terms, using, once again, the convenient function b_E.

Formally, a sequence of functions b_η is defined inductively on ordinals as follows: for any $P \subseteq W$,

(i) $b_0(P) = b_E(P)$,

(ii) If $\eta > 0$: $b_\eta(P) = b_E(\bigcap_{\zeta<\eta} b_\zeta(P) \cap P)$.

The sequence b_η can be seen to be decreasing: i.e. for all ordinals η, ζ and for all $P \subseteq W$, if $\eta < \zeta$ then $b_\zeta(P) \subseteq b_\eta(P)$. From the substitution schema, there is a smallest ordinal min such that for any $P \subseteq W$, $b_{\min+1}(P) = b_{\min}(P)$. The common belief operator b_C is then defined to be $b_{\min}$. It can be seen to satisfy the property that

$$\forall P \subseteq W,\ b_C(P) = b_E(b_C(P) \cap P)\,.$$

Let us denote by N_C the neighborhood function corresponding to b_C, i.e. $P \in N_C(w)$ iff $w \in b_C(P)$.

The validation clause for $C\psi$ is the usual one in a neighborhood structure:

$$\langle m, w\rangle \models C\psi \Leftrightarrow [\![\psi]\!]^m \in N_C(w)\,. \tag{$*$}$$

Clearly, $[\![C\psi]\!]^m = b_C([\![\psi]\!]^m)$. Using this semantics, Lismont (1993) proves that system MC_A is determined by $\mathcal{M}^N$.

The following theorem states that the validation clause (*) and the belief closure clause used throughout this paper are equivalent. It also states the important set-theoretic fact that the fixed-point construction underlying (*) *had* to be transfinite. In words, when it comes to neighborhood structures, validation clause (3) and the b.c. validation clause (2) are (nontrivially) equivalent, but clause (1), which is just based on a countably infinite iteration, is *not* equivalent to (2) or (3).

Theorem 6. (a) *Let $m = \langle W, (N_a)_{a\in A}, v\rangle$ be monotonic. Then, for any $w \in W$ and any $P \subseteq W$,*

$$P \in N_C(w) \Leftrightarrow \exists P' \in N_E(w)\,,\quad \textit{such that } P' \subseteq P \textit{ and } P' \textit{ is b.c.}$$

(b) *For any ordinal η there exists a neighborhood structure in which the ordinal* min *is equal to η.*

Proof of part (a). Equivalently, we have to prove that

$$\forall P \subseteq W,\ b_C(P) = \bigcup \{b_E(P') \mid P' \subseteq b_E(P') \cap P\}\,.$$

Taking $P' = b_C(P) \cap P$, the inclusion from the left to the right follows immediately from the fact that $b_C(P) = b_E(b_C(P) \cap P) = b_E(P')$.

To prove the converse inclusion, we inductively show that if $P' \subseteq b_E(P') \cap P$, then for any ordinal η, $b_E(P') \subseteq b_\eta(P)$.

(i) If $\eta = 0$: the monotonicity of b_E implies that $b_E(P') \subseteq b_E(P) = b_0(P)$,

(ii) If $\eta > 0$, from the inductive hypothesis:

$$P' \subseteq b_E(P') \cap P \subseteq \bigcap_{\zeta < \eta} b_\zeta(P) \cap P .$$

From the monotonicity of b_E, we conclude that

$$b_E(P') \subseteq b_E\Big(\bigcap_{\zeta < \eta} b_\zeta(P) \cap P\Big) = b_\eta(P) .$$

Proof of part (b). Define $W = \eta + 1$ and v in any way whatsoever. For any two ordinals $\zeta_1 \leq \zeta_2$, let us define $[\zeta_1, \zeta_2] = \{\xi \mid \zeta_1 \leq \xi \leq \zeta_2\}$. For all $a \in A$ and $P \subseteq W$, define:

(i) $b_a(P) = [\min(P) + 1, \eta]$ if P is nonempty and if $\min(P)$, i.e. the smallest ordinal in P, is less than η,

(ii) $b_a(P) = \emptyset$, otherwise.

Clearly, the functions b_a are monotonic, and for any $a \in A$, $b_E = b_a$. Now, we show that

$$\forall \zeta < \eta,\ b_\zeta(W) = [\zeta + 1, \eta] ,$$

using an inductive argument. First, if $\eta = 0$, $b_0(W) = b_E(W) = [1, \eta]$. Second, for any ordinal $\zeta < \eta$:

$$b_\zeta(W) = b_E\Big(\bigcap_{\xi < \zeta} b_\xi(W) \cap W\Big) = b_E\Big(\bigcap_{\xi < \zeta} [\xi + 1, \eta]\Big) = b_E([\zeta, \eta]) = [\zeta + 1, \eta] .$$

The following deals with the case $\zeta = \eta$:

$$b_\eta(W) = b_E\Big(\bigcap_{\xi < \eta} b_\xi(W) \cap W\Big) = b_E\Big(\bigcap_{\xi < \eta} [\xi + 1, \eta]\Big) = b_E([\eta, \eta]) = \emptyset .$$

This last result completes the proof that

$$\forall \zeta < \eta,\ b_{\zeta+1}(W) \neq b_\zeta(W) .$$

Hence $\min \geq \eta$.

To see the converse inequality, note that $b_{\eta+1}(W) = b_\eta(W) = \emptyset$, and therefore

$$\forall Q \subseteq W,\ b_{\eta+1}(Q) = b_\eta(Q) = \emptyset ,$$

from the properties of the b_ζ functions. Hence $\min \leq \eta$. □

Theorem 6 completes our attempt to mutually relate the fixed-point and iterate semantic accounts of common belief. We now discuss a technical problem that has attracted some attention among specialists of modal logic. This problem is unrelated to the main theme of the present section, but the above results and constructions happen to cast some light on its discussion. As emphasized in the previous section, completeness proofs for common belief systems are derived for each given formula ψ. This 'formula-by-formula' technique does not deliver *extended completeness*, i.e. the property that for any (finite or infinite) set of formulas Σ and any formula ψ:

$$\Sigma \models \psi \Rightarrow \Sigma \vdash \psi .$$

In the case of Normal axiom systems it is easy to demonstrate that extended completeness fails. Here is the (by now, well-known) counterexample.

Take $\Sigma = \{E^k p \mid k \geq 1\}$ and $\psi = Cp$, where p is any propositional variable. If m is any Kripke structure and w any world in it such that $\langle m, w \rangle \models \Sigma$, it is also the case that $\langle m, w \rangle \models Cp$. (Depending on the chosen definition of the validation clause for common belief, this implication holds either definitionally, or as a consequence of Proposition 5 above.) Hence, $\Sigma \models \psi$. Assume now that $\Sigma \vdash \psi$. Then, there would exist a *finite* subset $\Sigma' \subseteq \Sigma$ such that $\Sigma' \vdash \psi$ and (from the soundness theorem) $\Sigma' \models \psi$. But it is not difficult to construct a Kripke structure m' and a world w' in it such that $\langle m', w' \rangle \models \Sigma'$ and $\langle m', w' \rangle \models \psi$. (Specifically, the finiteness property of Σ' means that there is an upper limit K on the k such that $E^k p \in \Sigma'$; it is enough to make sure that $\langle m', w' \rangle \not\models E^{k+1} p$.)

Interestingly, the counterexample of the last paragraph does not say anything on extended completeness in the case of monotonic axiom systems. To see why this is so, take a monotonic neighborhood structure such as that exhibited in the proof of Theorem 6, part (b), choosing $\eta = \omega + 1$ and defining the valuation v in any way, provided that $[\![p]\!]^m = W$. For any $k \geq 1$, this model satisfies the property that $\langle m, w \rangle \models E^k p$ whenever $w \in [k, \omega + 1]$. Hence,

$$\langle m, \omega + 1 \rangle \models E^k p , \quad \forall k \geq 1 .$$

But $\langle m, \omega + 1 \rangle \not\models Cp$ follows from the fact mentioned in the proof that $b_C(W) = \emptyset$. Hence, the previous argument does not carry through.

More generally, we are not aware of any proof that extended completeness fails in the case of monotonic axiomatizations of common belief. Notice that if it existed, such a proof would also be a proof that monotonic axiomatizations are not *compact*, in the logician's usual sense of the word.

We end up this section by briefly discussing papers in economics and game theory that deliver anticipations of the belief closure approach to common belief, i.e. Milgrom (1981), Bacharach (1985), Monderer and Samet (1989), and Samet (1990).[10] Each of these four papers provides an independent restatement (or possibly weakening) of Aumann's (1976) classic definition of common knowledge

[10] For further references in this literature, see Lismont and Mongin (1994a). It would be interesting to sketch a comparison with alternative logical renderings of the fixed-point notions of common belief, e.g. Barwise (1989, Chapter 9), and Halpern and Moses (1990, Appendix A). Such a comparison would lead one far beyond the scope of this paper.

in terms of the meet of the agents' partitions. The two papers closest in spirit to our belief closure semantics are perhaps Milgrom (1981), and Monderer and Samet (1989). The former restates Aumann's definition by means of properties that are informally connected with (FP) and (RI) above (Milgrom, 1981, p. 220). The latter introduces the notion of an *evidently known event* as one that is known by everybody whenever it occurs. It then re-expresses common knowledge à la Aumann in the following way: an event E is common knowledge at a state iff there is an evidently known event E' such that E' occurs at the given state and E' implies that everybody knows E (see Monderer and Samet, 1989, pp. 174–175). Monderer and Samet's definition of evidently known events is identical to our definition of belief-closed subsets. Assuming monotonicity and the truth axiom, their definition of common knowledge events and our validation clause for $C\psi$ collapse into each other. But, in general, the two definitions are distinct. Heifetz (1994b) further elaborates on this point.

Another relevant game-theoretic connection is the hierarchical construction by Mertens and Zamir (1985). Roughly speaking, these authors exhibit a set W of worlds endowed with much internal structure: any $w \in W$ is an infinite sequence $\langle \theta, (P_a^1)_{a\in A}, (P_a^2)_{a\in A}, \ldots, \rangle$, where $\theta \in \Theta$ is the value of an objective parameter, P_a^1 is player a's subjective probability on Θ, P_a^2 is player a's subjective probability on the agents' subjective probabilities on Θ, and so on ad infinitum. These (countably) infinite sequences are constrained to satisfy coherence conditions that turn out to be interpretable epistemically. Mertens and Zamir's isomorphism theorem implies that each world w can be paired in a one-to-one way with a vector $\langle \theta, (\theta_a(w))_{a\in A} \rangle$ where the $\theta_a(w)$ are *subjective probabilities on W*. They investigate subsets P of W – 'belief subspaces' – with the following property:

$$\forall w \in P\,, \quad \forall a \in A,\ \theta_a(w)(P) = 1\,.$$

In words, a belief subspace is an event of the world space that is believed (in the sense of having probability 1) by everybody in every world in which it occurs. This notion is the authors' main tool to investigate common belief. There are visible analogies between the Mertens–Zamir approach and neighborhood semantics; Lismont (1992) provides further clarification. Note also the curious analogy between the above probabilistic hierarchy and the semantic construction of knowledge structures in Fagin et al. (1991), and of belief worlds in Vardi (1986); Mongin (1995) sketches a comparison with the latter semantics.

A common feature of the above-mentioned papers is that they have no syntactical component. To the epistemic logician they count mainly as preliminary clarifications of the various semantics that the formal operators may receive. An instructive conclusion from this literature is that the circular or fixed-point properties of common belief can be expressed easily in the semantics, which then becomes conveniently close to the (inevitably) circular syntax. This conclusion was the heuristic starting point for the logical elaborations provided in the present paper.

Acknowledgements

The authors are grateful to A. Heifetz and M. Kaneko for comments on earlier versions. They have also benefited from discussions with M. Bacharach, J. Halpern, T. Lucas, A. Rustichini, D. Samet, and the participants in the Epistemic Logic conferences in Liège (March 1993) and Marseille (January 1994). The authors acknowledge support from the Union Européenne programmes, 'Human Capital and Mobility', and SPES, respectively.

References

R.J. Aumann, Agreeing to disagree, Ann. Stat. 4 (1976) 1236–1239.

M. Bacharach, Some extensions of a claim of Aumann in an axiomatic model of knowledge, J. Econ. Theory 37 (1985) 167–190.

J. Barwise, The Situation in Logic, Centre for the Study of Language and Information, Stanford, 1989.

G. Boolos, The Unprovability of Consistency (Cambridge University Press, Cambridge, 1979).

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

M.J. Cresswell, The completeness of KW and K1.1, Logique et Analyse, no. 102 (1983) 123–127.

R. Fagin, J.Y. Halpern, Y.O. Moses and M.Y. Vardi, Reasoning About Knowledge (MIT Press, Cambridge, Mass., in press).

R. Fagin, J.Y. Halpern and M.Y. Vardi, A model-theoretic analysis of knowledge, J. Assoc. Comput. Mach. 38 (1991) 382–428. A preliminary version appeared in the Proc. 25th IEEE Symp. Found. Comput. Sci. (1984) 268–278.

J.D. Geanakoplos, Common knowledge, J. Econ. Persp. 6 (1992) 53–82.

R. Goldblatt, Logics of Time and Computation, Centre for the Study of Language and Information, Stanford, 1987.

J.Y. Halpern and Y.O. Moses, Knowledge and common knowledge in a distributed environment, J. Assoc. Comput. Mach. 37 (1990) 549–587. A preliminary version appeared in the Proc. 3rd ACM Conf. Princ. Distrib. Comput., New York (1984) 50–61.

J.Y. Halpern and Y.O. Moses, A guide to completeness and complexity for modal logics of knowledge and belief, Artif. Intell. 54 (1992) 319–379.

D. Harel, Dynamic logic, in: D. Gabbay and F. Guenther, eds., Handbook of Philosophical Logic, Vol. 2 (Reidel Publishing Company, Dordrecht, 1984) 497–604.

A. Heifetz, Infinitary epistemic logic, Mimeo, School of Mathematical Sciences, Tel Aviv University, 1994a.

A. Heifetz, Common belief in monotonic epistemic logic, Mimeo, School of Mathematical Sciences, Tel Aviv University, 1994b.

M. Kaneko and T. Nagashima, Final decisions, the Nash equilibrium and solvability in games with common knowledge of logical abilities, Math. Soc. Sci. 22 (1991) 229–255.

M. Kaneko and T. Nagashima, Indefinability of the common knowledge concept in finitary logics, Discussion Paper no. 551, Institute of Socio-Economic Planning, University of Tsukaba, 1993.

S. Kraus and D. Lehmann, Knowledge, belief and time, Theor. Comput. Sci. 58 (1988) 155–174.

L. Lismont, La connaisance commune: Approaches modale, ensembliste et probabiliste, Dissertation doctorale, Université Catholique de Louvain, 1992.

L. Lismont, La connaissance commune en logique modale, Math. Logic Q. 39 (1993) 115–130.

L. Lismont and P. Mongin, La connaissance commune: Une sémantique pour la logique modale, Logique et Analyse, no. 133–134 (1991) 133–149.

L. Lismont and P. Mongin, On the logic of common belief and common knowledge, Theory Dec. 37 (1994a) 75–106.

L. Lismont and P. Mongin, A non-minimal but very weak axiomatization of common belief, Art. Intell. 70 (1994b) 363–374.

J-F. Mertens and S. Zamir, Formulation of Bayesian analysis for games with incomplete information, Int. J. Game Theory 14 (1985) 1–29.

P. Milgrom, An axiomatic characterization of common knowledge, Econometrica 49 (1981) 219–222.

S. Modica and A. Rustichini, Awareness and partitional information structures, Theory Dec. 37 (1994) 107–124.

D. Monderer and D. Samet, Approximating common knowledge with common beliefs, Games Econ. Behav. 1 (1989) 170–190.

P. Mongin, Some connections between epistemic logic and the theory of nonadditive probability, in: P. Humphreys, ed., Patrick Suppes, Scientific Philosopher, Vol. 1 (Kluwer, Dordrecht, 1994), 135–172.

P. Mongin, Une interprétation logique du théorème de Mertens et Zamir, in: J. Dubucs and F. Lepage, ed., Méthodes logiques pour les sciences cognitives (Hermès, Paris, 1995) 235–259.

D. Samet, Ignoring ignorance and agreeing to disagree, J. Econ. Theory 52 (1990) 190–207.

M.Y. Vardi, On epistemic logic and logical omniscience, in: J.Y. Halpern, ed., Proc. 1st Conf. Theor. Aspects Reasoning about Knowledge (Morgan Kaufmann Publishers, Los Altos, 1986) 293–305.

M.Y. Vardi, On the complexity of epistemic reasoning, in: Proc. 4th Annual Symp. Logic Comput. Sci., 243–252.