Justified and Common Knowledge: Limited Conservativity

Evangelia Antonakos

CUNY Graduate Center
Ph.D. Program in Mathematics
365 Fifth Avenue
New York, NY 10016, USA
Eva@Antonakos.net

Abstract. We consider the relative strengths of three formal approaches to public knowledge: "any fool" knowledge by McCarthy (1970), Common Knowledge by Halpern and Moses (1990), and Justified Knowledge by Artemov (2004). Specifically, we show that epistemic systems with the Common Knowledge modality C are conservative with respect to Justified Knowledge systems on formulas $\chi \wedge C\varphi \rightarrow \psi$, where χ, φ, and ψ are C-free.

Keywords: justified knowledge, common knowledge, Artemov, conservative.

1 Multi-agent Logics

The logics T_n, $\mathsf{S4}_n$, and $\mathsf{S5}_n$ are logics in which each of the finitely many (n) agents has a knowledge operator K_i which is T, or $\mathsf{S4}$, or $\mathsf{S5}$ respectively. We only consider cases where all agents' modalities are of the same logical strength.

Definition 1. *The formal systems for* T_n, $\mathsf{S4}_n$, *and* $\mathsf{S5}_n$ *are as follows:*
Propositional Logic *plus for* K_i, $i = 1, 2, \ldots, n$ *we have*
Axioms for $\mathsf{S4}_n$**:**

$\mathsf{K} : K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi)$	each agent can do *modus ponens*
$\mathsf{T} : K_i\varphi \rightarrow \varphi$	agents can know only true propositions
$\mathsf{4} : K_i\varphi \rightarrow K_iK_i\varphi$	agents have positive introspection

Rules:

Necessitation: $\vdash \varphi \Rightarrow \vdash K_i\varphi$, *for* $i = 1, 2, \ldots, n$

For T_n, *omit the final axiom.*
For $\mathsf{S5}_n$, *add negative introspection:* $\neg K_i\varphi \rightarrow K_i\neg K_i\varphi$.

Definition 2. *Kripke models for* $\mathsf{S4}_n$*:* $M = \langle W,\ R_1,\ R_2, \ldots,\ R_n,\ \Vdash\rangle$ *where*

- W *is a non-empty set of worlds*
- $R_i \subseteq W \times W$ *is agent i's accessability relation.* R_i *is reflexive and transitive.*

S. Artemov and A. Nerode (Eds.): LFCS 2007, LNCS 4514, pp. 1–11, 2007.

• $\Vdash \subseteq W \times Var$ *where* Var *is the set of propositional variables. The forcing relation* $\Vdash$ *is naturally extended to all formulas so that* R_i *corresponds to* K_i*:*

$$M, u \Vdash K_i\varphi \Leftrightarrow (\forall v \in M)[uR_iv \rightarrow M, v \Vdash \varphi] \ .$$

For T_n-models, each R_i is reflexive while for $\mathsf{S5}_n$-models, each R_i is an equivalence relation.

Theorem 1. T_n*,* $\mathsf{S4}_n$*, and* $\mathsf{S5}_n$ *are sound and complete with respect to their models (cf.* [10]*).*

Multi-agent systems are enhanced by the addition of modalities which take into account shared or public knowledge of agents. Three such modalities C, J, and O will be discussed, all of which model variations of public information. We will compare their logical strengths, semantics, and complexity and will see why Justified Knowledge (J) systems are sufficient to solve classical epistemic scenarios, a role usually designated for Common Knowledge (C).

2 Common Knowledge

The most recognized concept of public knowledge is common knowledge, and the literature addressing it, both philosophical and mathematical, is vast. The initial investigation was philosophical: Lewis's book [15] on convention. The intuition behind the informal definition of common knowledge below derives from Aumann's oft-cited [5], where it was used in the context of agents having common priors. McCarthy's 'any fool' operator of 1970 ([10], p. 13) is closely related to common knowledge and his systems in [16] (see section 4 of this paper) may be the first to address it axiomatically . Rigorous work on common knowledge in the context of multi-agent systems was done by Halpern and Moses in [13] (an expansion of a 1984 work of the same title) and Lehmann [14]. Much of the work by Halpern and Moses appears in [10]. Common knowledge continues to be actively investigated.

Informally, the epistemic operator $C\varphi$, to be read 'φ is common knowledge,' can be given as infinite conjunction:

$$C\varphi \leftrightarrow \varphi \wedge E\varphi \wedge EE\varphi \wedge EEE\varphi \wedge E^4\varphi \wedge \cdots \wedge E^n\varphi \wedge \cdots \tag{1}$$

where $E\varphi = K_1\varphi \wedge K_2\varphi \wedge \cdots \wedge K_n\varphi$ ('everyone knows φ') and K_i is an individual agent's knowledge operator corresponding to T, $\mathsf{S4}$ or $\mathsf{S5}$ as appropriate. One formal characterization which [10] and [7] take is via the Fixed Point Axiom

$$C\varphi \leftrightarrow E\,(\varphi \wedge C\varphi) \tag{2}$$

and the Induction Rule

$$\frac{\varphi \rightarrow E\,(\varphi \wedge \psi)}{\varphi \rightarrow C\psi} \tag{3}$$

yielding Common Knowledge to be the greatest fixed point solution to $X \leftrightarrow E(\varphi \wedge X)$ [10]. Common Knowledge does not take into account the means by which the knowledge is acquired. As we will see, this is in contrast to Justified Knowledge. The distinction between the infinite conjunction, the fixed point axiom, and how common knowledge is achieved is addressed in [6]. [12] too, provides a survey with examples but does not include a distinct formalism. There is also an equivalent axiomatic formulation of common knowledge which replaces the induction rule with the induction axiom in [17], which, for technical convenience, we will use.

Definition 3. *T_n^C, $\mathsf{S4}_n^C$, and $\mathsf{S5}_n^C$ axiom systems:*
Propositional Logic *plus*
Axioms:

T, $\mathsf{S4}$, *or* $\mathsf{S5}$ *axioms for* K_i, $i = 1, 2, \ldots, n$, *respectively;*
K: $C(\varphi \to \psi) \to (C\varphi \to C\psi)$;
T: $C\varphi \to \varphi$;
$C\varphi \to E(C\varphi)$, *where* $E\varphi = K_1\varphi \wedge K_2\varphi \wedge \ldots \wedge K_n\varphi$;
Induction Axiom: $\varphi \wedge C(\varphi \to E\varphi) \to C\varphi$.

Rules:

Necessitation: $\vdash \varphi \Rightarrow \vdash K_i\varphi$, *for* $i = 1, 2, \ldots, n$
Necessitation: $\vdash \varphi \Rightarrow \vdash C\varphi$.

Definition 4. *Models for T_n^C, $\mathsf{S4}_n^C$, and $\mathsf{S5}_n^C$: $M = \langle W, R_1, R_2, \ldots, R_n, R_C, \Vdash\rangle$ where*

- $M = \langle W, R_1, R_2, \ldots, R_n, \Vdash\rangle$ *is a* T_n, $\mathsf{S4}_n$, *or* $\mathsf{S5}_n$ *model, respectively*
- $R_C = (\bigcup_{i=1}^{n} R_i)^*$, *that is the transitive closure of all the agents' relations*
- *The forcing relation* $\Vdash$ *is extended to all formulas so* R_C *corresponds to* C*:*

$$M, u \Vdash C\varphi \Leftrightarrow (\forall v \in M)[uR_Cv \to M, v \Vdash \varphi] \ .$$

Theorem 2. *T_n^C, $\mathsf{S4}_n^C$, and $\mathsf{S5}_n^C$ are sound and complete with respect to their models (cf.* [10], p. 70ff, [17], p. 47ff*).*

The agents' logic plays a role in determining the strength of the common knowledge operator C. In the systems defined above, C is always at least as strong as K_i. Showing that in T_n^C, $\mathsf{S4}_n^C$, and $\mathsf{S5}_n^C$, C satisfies the T, $\mathsf{S4}$, and $\mathsf{S5}$ axioms, respectively, is given as an exercise in [10], p. 93.

3 Justified Knowledge

Justified Knowledge was introduced by Artemov in [3,4] as the forgetful projection of the *evidence-based* knowledge represented by an appropriate adaptation of LP (Logic of Proofs). In LP systems ($\mathsf{T}_n\mathsf{LP}$, $\mathsf{S4}_n\mathsf{LP}$, $\mathsf{S5}_n\mathsf{LP}$), each formula / subformula carries with it a proof term representing a particular proof of the formula / subformula from the axioms. Justified knowledge systems are ones in which all proofs are identified as one. Whereas $C\varphi$ asserts that φ is common knowledge, $J\varphi$ asserts that φ is common knowledge arising from a proof of φ or some other agreed-upon acceptable set of evidences. Though the proof of φ is not explicitly presented with the assertion $J\varphi$, it is reproducible. This is the important Realization Theorem which provides an algorithm to reconstruct LP proof terms. For more details on this, the reader should consult [4].

As with the common knowledge logics, the construction of the justified knowledge logics T_n^J, $\mathsf{S4}_n^J$, and $\mathsf{S5}_n^J$ builds on the multi-agent logics. In C systems the agents' logic determines the strength of C while in J systems the strength of J is chosen independently to be weaker, stronger, or the same as that of the agents'. In the aforementioned logics, the modality J will be assumed to be $\mathsf{S4}$ unless otherwise specified.

Definition 5. T_n^J, $\mathsf{S4}_n^J$, *and* $\mathsf{S5}_n^J$ *axiom systems:*
Propositional Logic *plus*
Axioms:

T, $\mathsf{S4}$, *or* $\mathsf{S5}$ *axioms for* K_i, $i = 1, 2, \ldots, n$;
$\mathsf{S4}$ *axioms for* J ;
Connection Principle: $J\varphi \rightarrow K_i\varphi$.

Rules:

Necessitation for all K_i: $\vdash \varphi \Rightarrow \vdash K_i\varphi$;
Necessitation for J: $\vdash \varphi \Rightarrow \vdash J\varphi$.

Definition 6. *Models for* T_n^J, $\mathsf{S4}_n^J$, *and* $\mathsf{S5}_n^J$: $M = \langle W, R_1, R_2, \ldots, R_n, R_J, \Vdash \rangle$ *where*

- $M = \langle W, R_1, R_2, \ldots, R_n, \Vdash \rangle$ *is a* T_n, $\mathsf{S4}_n$, *or* $\mathsf{S5}_n$ *model, respectively*
- $R_J \subseteq W \times W$ *is reflexive and transitive relation such that* $R_J \supseteq (\bigcup_{i=1}^{n} R_i)^*$ *(where* * *is transitive closure)*
- *The forcing relation* $\Vdash$ *is extended to all formulas so* R_J *corresponds to* J:

$$M, u \Vdash J\varphi \Leftrightarrow (\forall v \in M)[uR_Jv \rightarrow M, v \Vdash \varphi] \ .$$

Theorem 3. T_n^J, $\mathsf{S4}_n^J$, *and* $\mathsf{S5}_n^J$ *are sound and complete with respect to their models, as shown in* [4].

Recall that in common knowledge models, $R_C = (\bigcup_{i=1}^{n} R_i)^*$ and so $R_C \subseteq R_J$. Thus in a context where we can compare the two, i.e. a hybrid model with both R_C and R_J, it seems (if φ contains no Js) $J\varphi \Rightarrow C\varphi$ but not vice versa. More formally, we have the following proposition.

Definition 7. *Let* φ^* *be* φ *with each instance of a* J *replaced by a* C.

Proposition 1. $(\mathsf{S4}_n^J)^* \subset \mathsf{S4}_n^C$ *but* $(\mathsf{S4}_n^J)^* \neq \mathsf{S4}_n^C$.

Proof. It needs to be shown that the $*$-translation of each each rule and axiom of $\mathsf{S4}_n^J$ is provable in $\mathsf{S4}_n^C$. Artemov shows this in [4] using the equivalent axiomatization of $\mathsf{S4}_n^C$ from [10]. It is only the Induction Axiom of $\mathsf{S4}_n^C$ which is not provable in $(\mathsf{S4}_n^J)^*$, yielding the strict inclusion. $\square$

The case in which the J of $\mathsf{S5}_n^J$ is an $\mathsf{S5}$ modality, that is $\mathsf{S5}_n^{J(\mathsf{S5})}$, is considered in [18] (where she names it $\mathsf{S5}_n\mathsf{S5}$). An $\mathsf{S5}_n^{J(\mathsf{S5})}$-model is like an $\mathsf{S5}_n^J$ model except that R_J will now be an equivalence relation.

Corollary 1. *Let* I.A. *be the induction axiom. Then*

$$\mathsf{S4}_n^C \equiv (\mathsf{S4}_n^J)^* + \mathsf{I.A.}\ ,$$
$$\mathsf{T}_n^C \equiv (\mathsf{T}_n^J)^* + \mathsf{I.A.}\ ,$$
$$\mathsf{S5}_n^C \equiv (\mathsf{S5}_n^{J(\mathsf{S5})})^* + \mathsf{I.A.}\ .$$

Proof. The strict inclusion of the J systems follows from Proposition 1 and noticing that C satisfies the $\mathsf{4}$ axiom in T_n^C and the $\mathsf{5}$ axiom in $\mathsf{S5}_n^C$. When the induction axiom is added, the equivalence is clear. $\square$

Indeed, from Corollary 1, in any of the justified knowledge systems mentioned, $J\varphi \Rightarrow C\varphi^*$.

The evidence-based common knowledge semantics for J systems are further enriched by Artemov's Realization Theorem mentioned at the start of the section. This gives a constructive approach to recovering or realizing the full proof terms of the evidence-based knowledge systems.

Theorem 4 (Realization Theorem). *There is an algorithm that, given an* $\mathsf{S4}_n^J$*-derivation of a formula* φ*, retrieves an* $\mathsf{S4}_n\mathsf{LP}$*-formula* ψ*, a* realization of φ*, such that* φ *is* ψ°*, where* $^\circ$ *replaces all proof terms with* J*, and* $\mathsf{S4}_n\mathsf{LP}$ *proves* ψ*.*

This theorem and a realization theorem for $\mathsf{S5}_n^J$ (where J is an $\mathsf{S4}$-modality) is established in [4] while a realization theorem for $\mathsf{S5}_n^{J(\mathsf{S5})}$ is given in [18].

Other major advantages to justified knowledge are

• proofs in $\mathsf{S4}_n^J$ are normalizable ([4]), but those in $\mathsf{S4}_n^C$ are not
• $\mathsf{S4}_2^J$ is *PSPACE*-complete [9], whereas for $n \geq 2$, $\mathsf{S4}_n^C$ is *EXPTIME*-complete [10].

These features have been exploited by Bryukhov in [8] to develop an automated theorem prover for $\mathsf{S4}_n^J$. Justified Knowledge offers simpler, more constructive, and more automation-friendly approach to common knowledge.

4 Any Fool's Knowledge

McCarthy's model of common knowledge via "any fool knows" apparently goes back to roughly 1970 ([10], p. 13), though its first published appearance is in [16]. In this epistemic multi-agent system, the modality for each agent is denoted by S, with an additional virtual agent, "any fool" denoted by O. In [16] p. 2, whatever any fool knows, "everyone knows that everyone else knows," and so someone knows. Thus we may add an additional axiom linking the fool to the other people: $O\varphi \rightarrow S\varphi$. Call this the linking axiom. This corresponds exactly to Artemov's connection principle: $J\varphi \rightarrow K_i\varphi$. When McCarthy et al. use subscripted modals, S_i, to specify individual agents, S_0 is the distinguished any fool operator O. Thus we see that the "fool" is a particular agent, hence in any axiom, we may replace all S modals by Os, though not vice versa.

Definition 8. *The McCarthy et al. axioms are based on propositional logic plus:*

linking axiom: $O\varphi \rightarrow S\varphi$
K0: $S\varphi \rightarrow \varphi$
K1: $O(S\varphi \rightarrow \varphi)$
K2: $O(O\varphi \rightarrow OS\varphi)$
K3: $O(S\varphi \wedge S(\varphi \rightarrow \psi) \rightarrow S\psi)$
K4: $O(S\varphi \rightarrow SS\varphi)$
K5: $O(\neg S\varphi \rightarrow S\neg S\varphi)$.

We will look at three systems identified in [16] given by axioms K0-K3, K0-K4, and K0-K5.[1] These will be referred to as MT, $\mathsf{M4}$, and $\mathsf{M5}$ respectively. Model semantics and completeness results for a variant of $\mathsf{M5}$ is stated in [16]. From this and Lemma 3 below, it follows that $\mathsf{S5}_n^{J(\mathsf{S5})}$ is also sound and complete.

These logics immediately lend themselves to epsitemic scenarios, of which Wise Men and Unfaithful Wives are addressed in [16]. These particular axioms do not seem to be built on standard formulations of modal logics yet we can see that Artemov's justified knowledge operator J plays a role equivalent to McCarthy's any fool operator O. In particular, we have the following theorem.

Definition 9. *Let $\varphi^\star$ be φ with each instance of a J replaced by an O and each K_i replaced by an S_i.*

Theorem 5. $(\mathsf{T}_n^J)^\star \equiv \mathsf{MT}$, $(\mathsf{S4}_n^J)^\star \equiv \mathsf{M4}$, *and* $(\mathsf{S5}_n^{J(\mathsf{S5})})^\star \equiv \mathsf{M5}$.

Proof. Immediate from the following three lemmas. □

Lemma 1. $(\mathsf{T}_n^J)^\star \equiv \mathsf{MT}$.

Proof. Recall that J is an $\mathsf{S4}$ modality while the K_i are T modalities.

($\Leftarrow$) To show $(\mathsf{T}_n^J)^\star \supset \mathsf{MT}$, T_n^J must satisfy MT axioms (K0-K3 and the linking axiom), where Os are Js and S_is are K_is.

Linking axiom: $\mathsf{T}_n^J \vdash J\varphi \rightarrow K_i\varphi$;	the connection principle
K0: $\mathsf{T}_n^J \vdash K_i\varphi \rightarrow \varphi$;	T axiom for K_i
K1: $\mathsf{T}_n^J \vdash J(K_i\varphi \rightarrow \varphi)$;	J necessitation of T axiom of K_i
K2: $\mathsf{T}_n^J \vdash J(J\varphi \rightarrow JK_i\varphi)$;	

[1] In [16], K0 is omitted from these lists. Given other statements in the paper, this clearly is just an oversight.

1. $\mathsf{T}_n^J \vdash J\varphi \to JJ\varphi$ — $\mathsf{4}$ axiom for J
2. $\mathsf{T}_n^J \vdash J\varphi \to K_i\varphi$ — the connection principle
3. $\mathsf{T}_n^J \vdash J(J\varphi \to K_i\varphi)$ — from 2. by J necessitation
4. $\mathsf{T}_n^J \vdash JJ\varphi \to JK_i\varphi$ — from 3. by K axiom for J
5. $\mathsf{T}_n^J \vdash J\varphi \to JK_i\varphi$ — from 1. and 4.
6. $\mathsf{T}_n^J \vdash J(J\varphi \to JK_i\varphi)$ — from 5. by J necesitation

K3: $\mathsf{T}_n^J \vdash J(K_i\varphi \wedge K_i(\varphi \to \psi) \to K_i\psi)$; — J necessitation of K axiom for K_i.

As mentioned above, “any fool” is a particular agent and so in any axiom all the Ss may be replaced by Os. Consider K0′-K3′ and the linking axiom′ where we do just that:

Linking axiom′: $\mathsf{T}_n^J \vdash J\varphi \to J\varphi$; — propositional tautology
K0′: $\mathsf{T}_n^J \vdash J\varphi \to \varphi$; — T axiom for J
K1′: $\mathsf{T}_n^J \vdash J(J\varphi \to \varphi)$; — J necessitation of T axiom of J
K2′: $\mathsf{T}_n^J \vdash J(J\varphi \to JJ\varphi)$; — J necessitaton of $\mathsf{4}$ axiom for J
K3′: $\mathsf{T}_n^J \vdash J(J\varphi \wedge J(\varphi \to \psi) \to J\psi)$; — J necessitation of K axiom for J.

($\Rightarrow$) $(\mathsf{T}_n^J)^\star \subset \mathsf{MT}$. We must show that MT satisfies the $(\mathsf{T}_n^J)^\star$ axioms and rules. Remember that “any fool” O is a particular S agent.

S axioms:
K: $\mathsf{MT} \vdash S\varphi \wedge S(\varphi \to \psi) \to S\psi$; — by K3, linking axiom, K0
T: $\mathsf{MT} \vdash S\varphi \to \varphi$; — K0

O axioms:
T: $\mathsf{MT} \vdash O\varphi \to \varphi$; — by K0, O is a particular S
K: $\mathsf{MT} \vdash O\varphi \wedge O(\varphi \to \psi) \to O\psi$; — by K axiom for S, O is a an S
4: $\mathsf{MT} \vdash O\varphi \to OO\varphi$; — by K2, T axiom for O, O is an S

Connection axiom: $\mathsf{MT} \vdash O\varphi \to S\varphi$; — the linking axiom

O necessitation: This follows from the fact that each S and O axiom is nesessitated.
K axiom for S and O is necessitated by K3.
T axiom for S and O is necessitated by K1.
$\mathsf{4}$ axiom for O is necessitated by K2.
S necessitation: This follows from O necessitation and the linking axiom. □

Lemma 2. $(\mathsf{S4}_n^J)^\star \equiv \mathsf{M4}$.

Proof. Recall that J and K_i are $\mathsf{S4}$ modalities.
($\Leftarrow$) $(\mathsf{S4}_n^J)^\star \supset \mathsf{M4}$ follows from Lemma 1 and
K4: $\mathsf{S4}_n^J \vdash J(K_i\varphi \to K_iK_i\varphi)$; — by J necessitation of $\mathsf{4}$ axiom for K_i
K4′ = K2′

($\Rightarrow$) $(\mathsf{S4}_n^J)^\star \subset \mathsf{M4}$ follows from Lemma 1 and
S axioms:
4: $\mathsf{M4} \vdash S\varphi \to SS\varphi$; — by K4, T axiom for O.
O necessitation: $\mathsf{4}$ axiom for S is necessitated by K4. □

Lemma 3. $(\mathsf{S5}_n^{J(S5)})^\star \equiv \mathsf{M5}$.

Proof. Recall that J and K_i are $\mathsf{S5}$ modalities.

($\Leftarrow$) $(\mathsf{S5}_n^{J(S5)})^\star \supset \mathsf{M5}$ follows from Lemma 2 and

K5: $\mathsf{S5}_n^{J(S5)} \vdash J(\neg K_i\varphi \to K_i\neg K_i\varphi)$; by J necessitation of $\mathsf{5}$ axiom for K_i.
K5′: $\mathsf{S5}_n^{J(S5)} \vdash J(\neg J\varphi \to J\neg J\varphi)$; by J necessitation of $\mathsf{5}$ axiom for J.

($\Rightarrow$) $(\mathsf{S5}_n^{J(S5)})^\star \subset \mathsf{M5}$ follows from Lemma 2 and

S axioms:

$\mathsf{5}$: $\mathsf{M5} \vdash \neg S\varphi \to S\neg S\varphi$; by K5, T axiom for O

O axioms:

$\mathsf{5}$: $\mathsf{M5} \vdash \neg O\varphi \to O\neg O\varphi$; by K5, T axiom for O, O is an S.

O necessitation: $\mathsf{5}$ axiom for O and S necessitated by K5. $\square$

Lemma 3 completes the proof of Theorem 5. Despite quite different motivations and technical backgrounds, McCarthy's "any fool" and Artemov's justified knowledge approaches lead to the same multi-modal logics.

Corollary 2. *There is a Realization Theorem for* MT*,* $\mathsf{M4}$*, and* $\mathsf{M5}$ *providing evidence-based semantics for McCarthy's "any fool" knowledge operator* O*.*

5 Limited Conservativity

A logic T with language $\mathcal{L}$ is a conservative extension of a logic T' with language $\mathcal{L}' \subseteq \mathcal{L}$ if for sentences φ of $\mathcal{L}'$, T proves φ if and only if T' proves φ. Recall the definition for * from Section 3 which renames J to C. As the logics $(\mathsf{S4}_n^J)^*$ and $\mathsf{S4}_n^C$ have the same language and yet are not equal, it is clear that $\mathsf{S4}_n^C$ can not be a conservative extension of $(\mathsf{S4}_n^J)^*$, it is however a conservative extension over all formulas in which C occurs only negatively. We say that a symbol or subformula X occurs negatively in a formula F if, when F is rewritten to have no implication symbols, X is in the scope of a negation symbol (or an odd number of negations). For example, X occurs only negatively in these first two formulas and both positively and negatively in the last: $X \to Y$, $(\neg(A \wedge X) \to B) \to Y$, $A \wedge X \to B \vee X$.

Theorem 6. *If* φ *is a formula of* $\mathsf{S4}_n^J$ *such that all occurrences of* J *in* φ *are negative, then* $\mathsf{S4}_n^J \vdash \varphi \quad \Leftrightarrow \quad \mathsf{S4}_n^C \vdash (\varphi)^*$.

In some sense this result is tight as the induction axiom $(\varphi \wedge J(\varphi \to E\varphi) \to J\varphi)$ which distinguishes $(\mathsf{S4}_n^J)^*$ from $\mathsf{S4}_n^C$ has, along with a negative occurrence of J, a single positive occurrence of J.

Proof. ($\Rightarrow$) is secured by the inclusion $(\mathsf{S4}_n^J)^* \subset \mathsf{S4}_n^C$ of Proposition 1.

($\Leftarrow$) This direction is a consequence of the Main Lemma which follows. We show this direction by proving the contrapositive. Suppose φ is a formula of $\mathsf{S4}_n^J$ such that all occurrences of J in φ are negative and $\mathsf{S4}_n^J \not\vdash \varphi$. By completeness, there is a model M and a world x such that $M, x \Vdash \neg\varphi$. By the Main Lemma, $M^C, x \Vdash^C \neg(\varphi)^*$, hence $\mathsf{S4}_n^C \not\vdash (\varphi)^*$, since M^C (with R_J ignored) is a model for $\mathsf{S4}_n^C$. □

Lemma 4 (Main Lemma). *Let M be a $\mathsf{S4}_n^J$-model. Add the relation R_C of reachability along $R_1, \ldots, R_n$ to M and get the augmented model M^C, where $\Vdash^C$ coincides with $\Vdash$ on variables, and the modality C corresponds to R_C. Let φ be a formula of $\mathsf{S4}_n^J$. Then*
if all occurrences of J in φ are positive, then $x \Vdash \varphi \quad \Rightarrow \quad x \Vdash^C (\varphi)^*$;
if all occurrences of J in φ are negative, then $x \not\Vdash \varphi \quad \Rightarrow \quad x \not\Vdash^C (\varphi)^*$.

Proof. By induction on φ.

Base case is secured by the definition of $\Vdash^C$.

Boolean case: $\varphi \equiv \psi \to \theta$.
Subcase: all occurrences of J in φ are positive and $x \Vdash \varphi$. Then $x \not\Vdash \psi$ or $x \Vdash \theta$. In the former case all occurrences of J in ψ are negative and, by the induction hypothesis, $x \not\Vdash^C (\psi)^*$. In the latter case all occurrences of J in θ are positive and, by the induction hypothesis, $x \Vdash^C (\theta)^*$. In either case, $x \Vdash^C (\varphi)^*$.
Subcase: all occurrences of J in φ are negative and $x \not\Vdash \varphi$. Then $x \Vdash \psi$ and $x \not\Vdash \theta$. Since all occurrences of J in ψ are positive and all occurrences of J in θ are negative, by the induction hypothesis, $x \Vdash^C (\psi)^*$ and $x \not\Vdash^C (\theta)^*$, hence $x \not\Vdash^C (\varphi)^*$.

Case: $\varphi \equiv K_i\psi$.
Subcase: all occurrences of J in φ are positive and $x \Vdash \varphi$. Then all occurrences of J in ψ are positive and $y \Vdash \psi$, for all y such that xR_iy. By the induction hypothesis, $y \Vdash^C (\psi)^*$, for all y such that xR_iy, hence $x \Vdash^C (K_i\psi)^*$, i.e., $x \Vdash^C (\varphi)^*$.
Subcase: all occurrences of J in φ are negative and $x \not\Vdash \varphi$. Then for some y such that xR_iy, $y \not\Vdash \psi$. Since all occurrences of J in ψ are also negative, by the induction hypothesis, $y \not\Vdash^C (\psi)^*$, hence $x \not\Vdash^C (K_i\psi)^*$, i.e., $x \not\Vdash^C (\varphi)^*$.

Case: $\varphi \equiv J\psi$.
Subcase: all occurrences of J in φ are positive and $x \Vdash \varphi$. Then all occurrences of J in ψ are also positive and $y \Vdash \psi$, for all y such that xR_Jy. Since $R_C \subseteq R_J$, $y \Vdash \psi$, for all y such that xR_Cy. By the induction hypothesis, $y \Vdash^C (\psi)^*$, for all y such that xR_Cy. Hence $x \Vdash^C C(\psi)^*$, i.e., $x \Vdash^C (J\psi)^*$, i.e., $x \Vdash^C (\varphi)^*$.
Subcase: 'all occurrences of J in φ are negative and $x \not\Vdash \varphi$' is impossible, since $\varphi \equiv J\psi$ and the displayed occurrence of J is positive in $J\psi$. □

Corollary 3. *If χ, φ and ψ are formulas in the language of $\mathsf{S4}_n$, then* $\mathsf{S4}_n^C \vdash \chi \wedge C\varphi \to \psi \Leftrightarrow \mathsf{S4}_n^J \vdash \chi \wedge J\varphi \to \psi$.

Proof. As per Theorem 6, $\chi \wedge J\varphi \to \psi$ has J only in negative position. □

Corollary 4. *If χ, φ and ψ are formulas in the language of T_n, then* $\mathsf{T}_n^C \vdash \chi \wedge C\varphi \to \psi \Leftrightarrow \mathsf{T}_n^J \vdash \chi \wedge J\varphi \to \psi$.

Proof. Analogous to the proof of Theorem 6 if the Main Lemma starts with T_n^J-models (J is an $\mathsf{S4}$ modality) and completeness for T_n^J. □

Corollary 5. *If χ, φ and ψ are formulas in the language of* $\mathsf{S5}_n$*, then* $\mathsf{S5}_n^C \vdash \chi \wedge C\varphi \rightarrow \psi \Leftrightarrow \mathsf{S5}_n^{J(\mathsf{S5})} \vdash \chi \wedge J\varphi \rightarrow \psi$.

Proof. Analogous to the proof of Theorem 6. if the Main Lemma starts with $\mathsf{S5}_n^{J(\mathsf{S5})}$-models ($J$ is an $\mathsf{S5}$ modality) and completeness for $\mathsf{S5}_n^{J(\mathsf{S5})}$. □

In fact, in Corollary 5, the justified knowledge system can be weakened to $\mathsf{S5}_n^J$, where J is an $\mathsf{S4}$ modality. Note that in the proof of Theorem 6, the case of $\varphi \equiv J\psi$ requires only that $R_C \subseteq R_J$ and not that R_J be of the same logical strength. The proof will actually go through with any modality whose accessibility relation contains R_C. However, if J is to be knowledge, R_J is semantically required to be reflexive and transitive. Thus for all formulas with only negative occurences of C, $\mathsf{S5}_n^C$ is a conservative extension of $(\mathsf{S5}_n^J)^*$.

6 Conclusions

This conservativity of C over J limited to formulas with C in negative position would seem to lend itself to uses of J in place of C in situations in which common knowledge is applied or assumed, rather than derived or concluded.

We have also seen that Artemov's evidence-based approach to common knowledge leads to the same multi-modal logic systems as McCarthy's 'any fool' axiomatic approach. This points towards applications of J and endows the O systems with a constructive, evidence-based semantics via the Realization Theorem.

We may also care to consider whether conservativity holds for a larger class of formulas and what benefits there may be to considering a logic which contains both J and C modalities.

Acknowledgements

Gratitude is due to Sergei Artemov for his guidance and for pointing me towards this problem. Thanks are due to Melvin Fitting for providing a preprint of his article referenced and to those who offered suggestions to drafts of this paper, including the reviewers of this volume.

References

1. L. Alberucci, G. Jaeger. About cut elimination for logics of common knowledge. *Annals of Pure and Applied Logic*, 133:73-99, 2005.
2. S. Artemov. Explicit provability and constructive semantics. *Bulletin of Symbolic Logic,* 7(1): 1-36, 2001.
3. S. Artemov. Evidence-Based Common Knowledge. Technical Report TR-2004018, CUNY Ph.D. Program in Computer Science, 2004.
4. S. Artemov. Justified Common Knowledge. *Theoretical Computer Science*, 357 (1-3): 4-22, 2006.
5. R. J. Aumann. Agreeing to Disagree. *Annals of Statistics*, 4(6): 1236-1239, 1976.
6. J. Barwise. Three Views of Common Knowledge. TARK 1988: 365-379.

7. J. van Benthem, D Sarenac. The Geometry of Knowledge. ILLC Report PP-2004-21, University of Amsterdam, 2004.
8. Y. Bryukhov. *Integration of decision procedures into high-order interactive provers.* Ph.D. thesis, CUNY Graduate School, 2005.
9. S. Demri. Complexity of Simple Dependent Bimodal Logics. Laboratoire LEIBNIZ-CNRS, U.M.R. 5522, manuscript, 2005.
10. R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning About Knowledge.* MIT Press, 1995.
11. M. Fitting. Modal Proof Theory. In *Handbook of Modal Logic*, Patrick Blackburn and Johan van Benthem and Frank Wolter, editors, Elsevier, 2006.
12. J. Geanakoplos. Common Knowledge. TARK 1992: 254-315.
13. J. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3): 549-587, 1990. (A preliminary version appeared in 1984.)
14. D. Lehmann. Knowledge, common knowledge, and related puzzles. *Proc.3rd ACM Symposium on Principles of Distributed Computing*, 62-67, 1984.
15. D. Lewis. *Convention, A Philosophical Study.* Harvard University Press, 1969.
16. J. McCarthy, M. Sato, T. Hayashi, and S. Igarishi. On the Model Theory of Knowledge. Technical Report STAN-CS-78-657, Stanford University, 1978.
17. J.-J. Ch. Meyer and W. van der Hoek. *Epistemic Logic of AI and Computer Science.* Cambridge Tracts in Theoretical Computer Science, 41. Cambridge University Press, 1995.
18. N. Rubtsova. On realization of $\mathsf{S5}$ modalitiy by evidence terms. *Journal of Logic and Computation*, 16: 671-684, 2006.