

Comparing Justified and Common Knowledge

Evangelia Antonakos

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

Abstract. What is public information in a multi-agent logic of knowledge such as T_n , $S4_n$ or $S5_n$? Traditionally this notion has been captured by *common knowledge*, which is an epistemic operator $C\varphi$ given roughly as $\bigwedge_{n=0}^{\infty} E^n\varphi$ where $E\varphi = K_1\varphi \wedge K_2\varphi \wedge \dots \wedge K_n\varphi$ (*everyone knows φ*) and K_i is an individual agent's knowledge operator. In common knowledge systems T_n^C , $S4_n^C$, and $S5_n^C$, C depends directly on the agents' logic ([1]). A stronger notion of public information is introduced in [3] with the new epistemic operator J . This provides *evidence-based common knowledge* for which $J\varphi$ (*φ is justified*) asserts that there is access to a proof of φ . Evidence-based systems are more flexible than their C counterparts as the logic for J can be chosen independently from that of the agents. The paper [3] considers evidence-based systems T_n^J , $S4_n^J$, and $S5_n^J$ obtained from the base logics T_n , $S4_n$, and $S5_n$ by adding a copy of **S4** for J and a connection principle $J\varphi \rightarrow K_i\varphi$, $i = 1, \dots, n$.

In this paper we compare common knowledge and justified knowledge in a canonical situation of the base logic **S4** and specify an exact logic principle that distinguishes them. Let I.A. stand for the "Induction Axiom"

$$\varphi \wedge J(\varphi \wedge E\varphi) \rightarrow J\varphi.$$

Theorem. $S4_n^C \equiv (S4_n^J + \text{I.A.})^*$ where $*$ is replacing J by C .
This theorem provides an alternative formulation for $S4_n^C$.

1 Introduction

The concept of common knowledge arises naturally in the context of a multi-modal logic of knowledge such as T_n , $S4_n$ or $S5_n$. These are logics of n -many rational agents, each with **T**, **S4** or **S5** reasoning, respectively. Common knowledge is an epistemic operator $C\varphi$ given roughly as an infinite conjunction

$$C\varphi \leftrightarrow \varphi \wedge E\varphi \wedge E^2\varphi \wedge E^3\varphi \wedge \dots \wedge E^n\varphi \dots$$

where $E\varphi = K_1\varphi \wedge K_2\varphi \wedge \dots \wedge K_n\varphi$ (*everyone knows φ*) and K_i is an individual agent's knowledge operator corresponding to **T**, **S4** or **S5** as appropriate. Traditionally, common knowledge has been characterised by the Fixed Point Axiom

$$C\varphi \leftrightarrow E(\varphi \wedge C\varphi)$$

and the Induction Rule

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

yielding common knowledge to be the greatest fixed point solution [1]. Being the most general solution to the Fixed Point Axiom, common knowledge does not take into account the means by which the knowledge is acquired.

Artemov has introduced a related notion of public information in [3] with the epistemic operator J which is an *evidence-based* common knowledge. 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. Read $J\varphi$ as φ is *justified*. This difference between common knowledge and knowledge based on evidence may be seen in the following informal example.

Aliki and Basil visit their a friend Christa who is very good at slight-of-hand. Christa sets up a card trick and Aliki selects a card from the deck and does not look at it or show it to either Basil or Christa. Christa declares that Aliki's card is the seven of clubs. Since Christa is skilled at card tricks, Aliki and Basil both believe, and say so, that Aliki's card is the seven of clubs. The card Aliki has is indeed the seven of clubs, so it is in fact common knowledge, not just belief, that this is her card. However it is only when Aliki actually turns over the card do they acquire evidence-based knowledge that this is the case.

In this paper we compare common knowledge and justified knowledge in a canonical situation of the base logic $S4$ and specify an exact logical principle that distinguishes them.

2 Multi-agent logics and common knowledge

A Kripke model for a multi-agent logic is a frame $(W, R_1, R_2, \dots, R_n)$, where W is a set of worlds and R_i is an individual agent's reachability relation on W , with a forcing relation \Vdash assigning propositional formulas to worlds. In models of T_n each R_i is reflexive, for $S4_n$ -models each R_i is reflexive and transitive, for $S5_n$ -models R_i s are reflexive, transitive, and symmetric. The modality K_i denotes agent i 's knowledge and we say $K_i\varphi$ if φ holds at each world in R_i . We will only consider cases in which all K_i are equivalent, in particular we will be concerned with extensions of $S4_n$. The axiom system for $S4_n$ is a copy of the $S4$ axioms for each modal K_i . For $S4_n^C$ we extend the axioms to include the Fixed Point Axiom and the Induction Rule for the common knowledge modal C as well as the definition of $E\varphi$ as $\bigwedge_{i=1}^n K_i\varphi$ [1]. A $S4_n^C$ -model is like an $S4_n$ model with an additional accessibility relation R_C on W for C which is the transitive closure of $\bigcup_{i=1}^n R_i$. The logic of C arises directly from the logic of the agents – if the K_i are changed from $S4$ to T then C changes accordingly.

3 Justified Knowledge

Evidence-based systems with the modality J are more flexible than their C counterparts as the logic for J can be chosen independently from that of the agents. \mathbb{T}_n^J , $\mathbb{S4}_n^J$, and $\mathbb{S5}_n^J$ can be axiomatized by adding a copy of $\mathbb{S4}$ for J and a connection principle

$$J\varphi \rightarrow K_i\varphi$$

for all $i = 1, \dots, n$ to the axiom systems of \mathbb{T}_n , $\mathbb{S4}_n$, and $\mathbb{S5}_n$. This principle is dubbed the "undeniability of evidence" principle in [3].

J can be viewed as an agent who accepts facts ($J\varphi$) when there is checkable evidence or a proof. J is like a repository of knowledge available to all agents. In mathematics, for example, let us consider a peer-reviewed proof to be our standard for knowledge. Then, if φ is some mathematical statement, $J\varphi$ holds just in case we have evidence for it, *i.e.* a peer-reviewed proof. In this case we say φ is a theorem. Each mathematician is justified in stating Fermat's Last Theorem is true, without having to supply the proof himself, because there is access to an accepted proof. decades ago, Fermat's Last Theorem may have been common knowledge but it was not evidence-based knowledge in this sense.

To build a model for these evidence-based systems start with a model of \mathbb{T}_n , $\mathbb{S4}_n$, and $\mathbb{S5}_n$ and add a reflexive transitive accessibility relation R_J on W for J such that $R_i \subseteq R_J$ for all $i = 1, \dots, n$. Note that

$$R_J \supseteq \text{transitive closure of } \bigcup_{i=1}^n R_i$$

but

$$R_C = \text{transitive closure of } \bigcup_{i=1}^n R_i.$$

See [3] for proofs that these logics are sound and complete with respect to these models.

Proposition 1 *Evidence-based knowledge J satisfies the Fixed-Point Axiom in each of \mathbb{T}_n^J , $\mathbb{S4}_n^J$, and $\mathbb{S5}_n^J$.*

Proof. We prove that $\mathbb{T}_n^J \vdash J\varphi \leftrightarrow E(\varphi \wedge J\varphi)$, and hence so do $\mathbb{S4}_n^J$ and $\mathbb{S5}_n^J$.
(\rightarrow)

1. $\mathbb{T}_n^J \vdash J\varphi \rightarrow K_i\varphi$, for $i = 1, 2, \dots, n$ by the connection axiom
2. $\mathbb{T}_n^J \vdash J\varphi \rightarrow E\varphi$ from 1. by definition of E
3. $\mathbb{T}_n^J \vdash J\varphi \rightarrow JJ\varphi$ 4 axiom (J is $\mathbb{S4}$)
4. $\mathbb{T}_n^J \vdash J\varphi \rightarrow EJ\varphi$ from 3. and 2.
5. $\mathbb{T}_n^J \vdash J\varphi \rightarrow E\varphi \wedge EJ\varphi$ from 2. and 4.
6. $\mathbb{T}_n^J \vdash J\varphi \rightarrow E(\varphi \wedge J\varphi)$

(\leftarrow)

1. $\mathbb{T}_n^J \vdash E(\varphi \wedge J\varphi) \rightarrow K_i(\varphi \wedge J\varphi)$ for $i = 1, 2, \dots, n$ by definition of E
2. $\mathbb{T}_n^J \vdash K_i(\varphi \wedge J\varphi) \rightarrow (\varphi \wedge J\varphi) \rightarrow J\varphi$ by the \mathbb{T} axiom for K_i
3. $\mathbb{T}_n^J \vdash E(\varphi \wedge J\varphi) \rightarrow J\varphi$

□

4 Comparing J and C

In order to compare J and C we must ensure there is model and a language which can support two logics, one with J and one with C . Consider the models of $\mathbb{S4}_n^C$ and $\mathbb{S4}_n^J$ based on the same $\mathbb{S4}_n$ model. A $\mathbb{S4}_n^C$ -model is an $\mathbb{S4}_n^J$ -model since the accessibility relation R of J is sure to include reachability on the frame (W, R_1, \dots, R_n) , which is the accessibility of relation of C . Though an $\mathbb{S4}_n^J$ -model is not an $\mathbb{S4}_n^C$ -models it is possible to augment it with an additional reachability relation for C . This beefed-up model supports the languages of both $\mathbb{S4}_n^C$ and $\mathbb{S4}_n^J$ ([3]). It is in the context of such a model that we will compare knowledge operators J and C . For instance, it is not hard to see that evidence-based knowledge is stronger than common knowledge, *i.e.* $J\varphi \rightarrow C\varphi$, but not vice versa ([3]).

omniscient; all consequences of the known formulas are also known. For example, if people were logically omniscient, anyone who knew the rules for chess would also know whether White has a winning strategy. Evidence-based systems are not, even when J is a more powerful logic than the agents' as in the case of \mathbb{T}_n [3].

In the following, let φ^* be φ with each instance of a J replaced by a C .

Proposition 2 *Each evidence-based principle is a common knowledge principle, *i.e.*, $(\mathbb{S4}_n^J)^* \subseteq \mathbb{S4}_n^C$, but not vice versa.*

Proof. We need to show that the *-translation of each each rule and axiom of $\mathbb{S4}_n^J$ is provable in $\mathbb{S4}_n^C$.

- Necessitation rule for J : $\mathbb{S4}_n^J \vdash \psi \Rightarrow \mathbb{S4}_n^J \vdash J\psi$.
Suppose $\mathbb{S4}_n^C \vdash \psi^*$, we want to show that $\mathbb{S4}_n^C \vdash C\psi^*$.
 1. $\mathbb{S4}_n^C \vdash \psi^*$ by assumption
 2. $\mathbb{S4}_n^C \vdash \psi^* \wedge \top$
 3. $\mathbb{S4}_n^C \vdash K_i(\psi^* \wedge \top)$ for $i = 1, 2, \dots, n$ by necessitation of K_i
 4. $\mathbb{S4}_n^C \vdash E(\psi^* \wedge \top)$ by definition of E
 5. $\mathbb{S4}_n^C \vdash \top \rightarrow E(\psi^* \wedge \top)$
 6. $\mathbb{S4}_n^C \vdash \top \rightarrow C\psi^*$ by induction rule of $\mathbb{S4}_n^C$
 7. $\mathbb{S4}_n^C \vdash C\psi^*$
- Modus Ponens rule involves no modals and is a rule of both logics.
- Connection axiom for J : $\mathbb{S4}_n^J \vdash J\varphi \rightarrow K_i\varphi$.
We want to show that $\mathbb{S4}_n^C \vdash C\varphi^* \rightarrow K_i\varphi^*$.
 1. $\mathbb{S4}_n^C \vdash C\varphi^* \leftrightarrow E(\varphi^* \wedge C\varphi^*)$ the fixed point axiom of $\mathbb{S4}_n^C$
 2. $\mathbb{S4}_n^C \vdash C\varphi^* \rightarrow E(\varphi^* \wedge C\varphi^*)$

3. $S4_n^C \vdash C\varphi^* \rightarrow K_i(\varphi^* \wedge C\varphi^*)$
4. $S4_n^C \vdash C\varphi^* \rightarrow K_i\varphi^* \wedge K_i(C\varphi^*)$
5. $S4_n^C \vdash C\varphi^* \rightarrow K_i\varphi^*$

– S4 axioms

It must be shown that the S4 axioms (K, T and 4) which hold for J , also hold for C , *i.e.* $S4_n^C \vdash C(\varphi \rightarrow \psi) \rightarrow (C\varphi \rightarrow C\psi)$, $S4_n^C \vdash C\varphi \rightarrow \varphi$, and $S4_n^C \vdash C\varphi \rightarrow CC\varphi$. These are given as exercises in [1] which is an excellent resource for common knowledge in logics of knowledge. The exercises are not complicated. For example, that the T axiom holds for C follows immediately from the fixed point axiom and induction rule for C .

To complete the proof the reverse inclusion must not hold, *i.e.* $(S4_n^J)^* \not\subseteq S4_n^C$. This is shown in [3] as follows. We want there to be a valid $S4_n^C$ principle whose translation in $S4_n^J$ is not valid. Consider the $S4_n^C$ principle $\iota_C = \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow C\varphi$. (That this is valid is shown in the proof of **lemma 1**.) In $S4_n^J$ this becomes $\iota_J = \varphi \wedge J(\varphi \rightarrow E\varphi) \rightarrow J\varphi$.

Artemov provides a two-world counter-model to this: Let $W = \{a, b\}$ and for all i only $bR_i b$ and $aR_i b$ (world b is accessible from world a). In addition, $bR_J a$. Let φ be such that $a \not\models \varphi$ but $b \models \varphi$. Then $b \models J(\varphi \rightarrow E\varphi)$ since $E\varphi$ holds everywhere φ does. Hence $b \models \varphi \wedge J(\varphi \rightarrow E\varphi)$. However, since $a \not\models \varphi$ and $bR_J a$, we have $b \not\models J\varphi$. Thus ι_J is not valid in $S4_n^J$ so the reverse inclusion fails to hold.

□

The principle ι_J from the end of the previous proof amounts to induction; let us add this as an induction axiom (I.A.) to $S4_n^J$. Call this expanded logic $S4_n^J + \text{I.A.}$. The I.A. principle not only distinguishes $S4_n^J$ from $S4_n^C$ but we will see that $S4_n^J + \text{I.A.}$ is equivalent to $S4_n^C$.

Lemma 1 *The induction axiom I.A. of $S4_n^J + \text{I.A.}$ holds in $S4_n^C$.*

(Henceforth the *s are omitted from proofs for easier reading.)

Proof. I.A.* = $\iota_C = \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow C\varphi$. We must show that $S4_n^C \vdash \iota_C$.

1. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow C(\varphi \rightarrow E\varphi)$ a tautology
2. $S4_n^C \vdash C(\varphi \rightarrow E\varphi) \rightarrow E(C(\varphi \rightarrow E\varphi))$ follows from the Fixed Point Axiom
3. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow E(C(\varphi \rightarrow E\varphi))$ from 1. and 2.
4. $S4_n^C \vdash C(\varphi \rightarrow E\varphi) \rightarrow (\varphi \rightarrow E\varphi)$ from the Fixed Point Axiom
5. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow E\varphi$ exportation
6. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow E\varphi \wedge E(C(\varphi \rightarrow E\varphi))$ from 3. and 5.
7. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow E(\varphi \wedge C(\varphi \rightarrow E\varphi))$
8. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow E(\varphi \wedge \varphi \wedge C(\varphi \rightarrow E\varphi))$
9. $S4_n^C \vdash \varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow C\varphi$ induction rule of $S4_n^C$

□

Proposition 3 $(S4_n^J + \text{I.A.})^* \subseteq S4_n^C$.

Proof. The proof follows immediately from **Proposition 2** and **Lemma 1**.

□

Lemma 2 *The induction rule for $S4_n^C$ holds in $S4_n^J + \text{I.A.}$.*

Proof. We need to show that $S4_n^J + \text{I.A.} \vdash \varphi \rightarrow E(\psi \wedge \varphi) \Rightarrow S4_n^J + \text{I.A.} \vdash \varphi \rightarrow J\psi$

1. $S4_n^J + \text{I.A.} \vdash \varphi \rightarrow E(\psi \wedge \varphi)$ by assumption
2. $S4_n^J + \text{I.A.} \vdash \varphi \rightarrow E\psi \wedge E\varphi$
3. $S4_n^J + \text{I.A.} \vdash \varphi \rightarrow E\varphi$
4. $S4_n^J + \text{I.A.} \vdash J(\varphi \rightarrow E\varphi)$ by J necessitation
5. $S4_n^J + \text{I.A.} \vdash \varphi \wedge J(\varphi \rightarrow E\varphi) \rightarrow J\varphi$ the induction axiom
6. $S4_n^J + \text{I.A.} \vdash J(\varphi \rightarrow E\varphi) \rightarrow (\varphi \rightarrow J\varphi)$ exportation
7. $S4_n^J + \text{I.A.} \vdash \varphi \rightarrow J\psi$ from 4. and 6. by modus ponens

□

Proposition 4 *Each common knowledge principle is an evidence-with-induction principle, i.e. $S4_n^C \subseteq (S4_n^J + \text{I.A.})^*$.*

Proof. We need to show that the *-translation of each rule and axiom of $S4_n^C$ is provable in $(S4_n^J + \text{I.A.})^*$.

- Fixed Point Axiom for C : $S4_n^C \vdash C\varphi \leftrightarrow E(\varphi \wedge C\varphi)$.
We want to show that $S4_n^J + \text{I.A.} \vdash J\varphi \leftrightarrow E(\varphi \wedge J\varphi)$. This is proved by **Proposition 1**.
- Induction Rule for C : $S4_n^C \vdash \varphi \rightarrow E(\psi \wedge \varphi) \Rightarrow S4_n^C \vdash \varphi \rightarrow C\psi$.
We want to show that $S4_n^J + \text{I.A.} \vdash \varphi \rightarrow E(\psi \wedge \varphi) \Rightarrow S4_n^J + \text{I.A.} \vdash \varphi \rightarrow J\psi$.
This is **Lemma 2**.

□

Comparing **Proposition 4.** and **Proposition 2.** we see that the induction axiom A.I. is the exact difference between $S4_n^C$ and $S4_n^J$.

Theorem 1 $(S4_n^J + \text{I.A.})^* \equiv S4_n^C$

Proof. **Proposition 3.** and **Proposition 4.**

5 Conclusion

We now have an alternative formulation of the familiar $S4_n^C$ in a setting which corresponds to a new evidence-based knowledge modality J .

I thank Sergei Artemov for introducing this material to me and for his assistance.

References

1. R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning About Knowledge*. MIT Press, 1995.
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.