Forms of Generic Common Knowledge

by

Evangelia Antonakos

A dissertation submitted to the Graduate Faculty in Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

2013

This manuscript has been read and accepted for the Graduate Faculty in Mathematics in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

Sergei Artemov

Date

Chair of Examining Committee

Linda Keen

Date

Executive Officer

Sergei Artemov

Melvin Fitting

Roman Kossak

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Forms of Generic Common Knowledge

by

Evangelia Antonakos

Advisor: Sergei Artemov

In multi-agent epistemic logics, common knowledge has been a central consideration of study. A generic common knowledge ($G.C.K.$) system is one that yields iterated knowledge $I(\varphi)$: 'any agent knows that any agent knows that any agent knows… φ' for any number of iterations. Generic common knowledge yields iterated knowledge

$$G.C.K.(\varphi) \rightarrow I(\varphi)$$

but is not necessarily logically equivalent to it. This contrasts with the most prevalent formulation of common knowledge C as equivalent to iterated knowledge. A spectrum of systems may satisfy the $G.C.K.$ condition, of which C is just one. It has been shown that in the usual epistemic scenarios, $G.C.K.$ can replace conventional common knowledge and Artemov has noted that such standard sources of common knowledge as public announcements of atomic sentences generally yield $G.C.K.$ rather than C.

In this dissertation we study mathematical properties of generic common knowledge and compare them to the traditional common knowledge notion. In particular, we contrast the modal $G.C.K.$ logics of McCarthy (e.g. $\mathsf{M4}$) and Artemov (e.g. $\mathsf{S4}_n^J$) with C-systems (e.g. $\mathsf{S4}_n^C$) and present a joint $C/G.C.K.$ implicit knowledge logic $\mathsf{S4}_n^{CJ}$ as a conservative extension of both. We show that in standard epistemic scenarios in which common knowledge of certain premises is assumed, whose conclusion does not concern common knowledge (such as Muddy Children, Wise Men, Unfaithful Wives, etc.), a lighter $G.C.K.$ can be used instead of the traditional, more complicated, common knowledge. We then present the first fully explicit $G.C.K.$ system $\mathsf{LP}_n(\mathsf{LP})$. This justification logic realizes the corresponding modal system $\mathsf{S4}_n^J$ so that $G.C.K.$, along with individual knowledge modalities, can always be made explicit.

Acknowledgements

My teacher and advisor Sergei Artemov has all my thanks and sincere gratitude. This dissertation would not exist with out his mentoring and indeed with out his research on which this work is based. He has guided me through this process with good advice, generous amounts of time and patience, financial support, and many engaging conversations. It has been stimulating to witness this burgeoning area of justification logic he planted, and to play a small part. Melvin Fitting is responsible for my enthusiasm for modal logic. His *First Order Modal Logic* written with Richard L. Mendelsohn was my constant companion for a year. His papers on justification logics have greatly enriched my understanding as has his deep knowledge of mathematical logic history. Roman Kossak encouraged me to apply to the Graduate Center, taught my Logic I and model theory classes, welcomed me to Bronx Community College when I began there as a teaching fellow (even giving me a private tour), and now has ushered me through my defense. I am most fortunate to have the support of these fine mathematicians and fine men. I thank these “Three Wise Men” for being my committee and for the many hats they have worn in my graduate student journey.

At the Graduate Center, I thank the Ph.D. Program in Mathematics and its EOs for financial support and its APO Rob Landsman, who knows all procedural details, for fostering such a collegial atmosphere. Colleagues from the Computational Logic Seminar, many my academic siblings and good friends, have been crucial: Yegor Bryukhov, Walter Dean, Hidenori Kurokawa, Roman Kuznets, Natalia Novak, Tudor Protopopescu, Bryan Renne, Çağil Taşdemir, Ren-June Wang, Junhua Yu.

My parents, Stephen and Naomi Spector Antonakos, are the best.

Over my years at the GC, many have made my intellectual and personal life more enjoyable and rewarding. A few such friends and faculty are: Karen Kletter for laughter and good cheeses; Shoshi Friedman for getting through our quals; Lily for insight; Rohit Parikh for his PDL and Game Theory classes; Jeanne Funk; Susan Schweitzer; Russell Miller; Stathis Zachos; Elena Nogina; Rehana Patel; Marjorie Senechal; Michael Carlisle; Leah Schnelbach; Dustin Mulcahey; Terence Kivran-Swaine; Marcos Zyman; Peggy Dean; Marianna Bonanome; Sheila Miller; Joel Hamkins; my BCC colleagues. Michael Schubert: Thank you for knowing I would finish. You made it easier.

Contents

Chapter 1

Introduction

Imbuing a formal logical theory with knowledge semantics has been a subject of intense study since the mid-20th century. Most of these epistemic logics have been modal logics, in which the formula $K\varphi$ is interpreted as 'φ is known' or $K_i\varphi$ as 'agent i knows φ.' The modality K is assigned axioms which reflect some approximation of the nature of knowledge, such as $K\varphi \to \varphi$, 'whatever is known, is true.'

The start of the 21st century has seen a related field of justification logic take root and flourish. Applying epistemic semantics to these logics gives formulas of the form $t{:}\varphi$ the reading 'φ is known due to justification t' or 't is a proof of φ.' Individual proof terms such as t are not themselves modals but collectively play a similar role. Accepting Plato's definition of knowledge as justified, true belief ([Pla]), these logics can be seen to address the missing component of justification absent from the modal logic approach, which is generally seen as modeling knowledge as true belief ([AF11]). We distinguish these two considerations of knowledge, referring to epistemic modal logics as systems of "implicit knowledge" and to justification logics as those of "explicit knowledge."

In systems of multiple knowers, or agents, it is natural to consider what information is publicly known. The most investigated such concept is that of common knowledge. Informally, if a sentence or proposition ϕ is common knowledge, $C\phi$, then everyone knows it, and everyone knows everyone knows it, and everyone knows everyone knows everyone knows it, etc., i.e., iterated knowledge of ϕ, $I\phi$. Common knowledge has overwhelmingly been formalized as an equivalence of $C\phi$ and $I\phi$ via a finite set of axioms; this ill-defined infinite conjunction of I is not essential. In each multi-agent system, C is unique. However, there is an alternate conception of common knowledge, generic common knowledge, $G.C.K.$ While $G.C.K.$ is sufficient to yield iterated knowledge, it is not necessarily equivalent to it. This approach offers a broader view of common knowledge as it allows for a choice between multiple logically non-equivalent common knowledge operators. Both notions have been used to analyze classic epistemic scenarios such as Muddy Children.

After a development of background material in this chapter, in the following chapter we compare C and $G.C.K.$ systems, $\mathsf{S4}_n^C$ and $\mathsf{S4}_n^J$, in the implicit setting, and present and explore a joint implicit $C/G.C.K.$system $\mathsf{S4}_n^{CJ}$. The third chapter will address generic common knowledge in a fully explicit setting for the first time, in $\mathsf{LP}_n(\mathsf{LP})$, and provide a Realization Theorem which shows this system to be a direct explicit counterpart to $\mathsf{S4}_n^J$.

1.1 Multi-Agent Modal Logics

We begin with an introduction to modal logics with epistemic semantics. Whereas in propositional modal logic $\Box\varphi$ is often read 'φ is necessary,' it is standard when modeling knowledge to use K as the modality and read $K\varphi$ as 'φ is known,' or 'φ is knowable.' We consider multi-agent systems as they are a natural context for common knowledge. The modal logics T_n, $\mathsf{S4}_n$, and $\mathsf{S5}_n$ are ones in which each of the finitely many (n) agents has a knowledge operator K_i which is a modality of the logic T, or $\mathsf{S4}$, or $\mathsf{S5}$ respectively. We will only consider cases where all agents' modalities are of the same logical strength.

Definition 1. The *language* $\mathcal{L}_{\mathsf{S4}_n}$ is an extension of the propositional language:

$$\mathcal{L}_{\mathsf{S4}} := \{\, Var, \wedge, \vee, \rightarrow, \neg, K_i\}$$

for $i \in \{1, 2, \dots, n\}$ where Var is the set of propositional variables. *Formulas* are defined by the grammar

$$\varphi := p \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi \mid \neg\varphi \mid K_i\varphi$$

where $p \in \mathit{Var}$.

The formula $K_1K_2\varphi$ has the intended semantics of 'agent 1 knows that agent 2 knows φ.' The languages and formulas of T_n and $\mathsf{S5}_n$ are those of $\mathsf{S4}_n$.

Definition 2. The *axioms and rules of* $\mathsf{S4}_n$:

CLASSICAL PROPOSITIONAL LOGIC:

A. axioms of classical propositional logic

R1. $\vdash (\varphi \to \psi) \wedge \varphi \Rightarrow\ \vdash \psi$ *modus ponens*

$\mathsf{S4}$-KNOWLEDGE AXIOMS FOR ALL AGENTS $K_i,\ i \in \{1, 2, \ldots, n\}$:

K. $K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$ *each agent can do modus ponens*

T. $K_i\varphi \to \varphi$ *agents can know only true propositions*

$\mathsf{4}$. $K_i\varphi \to K_iK_i\varphi$ *agents have positive introspection*

MODAL NECESSITATION RULE FOR ALL AGENTS:

R2. $\vdash \varphi \Rightarrow\ \vdash K_i\varphi$.

For the logic T_n, omit axiom $\mathsf{4}$. For $\mathsf{S5}_n$, add the negative introspection axiom $\mathsf{5}$: $\neg K_i\varphi \to K_i\neg K_i\varphi$.

Note that while the axioms of these logics do not involve more than a single modality, their formulas may. In particular, repeated uses of R2 to the axioms yields that each agent knows that the others are also $\mathsf{S4}$-type agents, e.g. $K_2(K_1\varphi \to K_1K_1\varphi)$. Multi-agent systems are enhanced by the addition of modalities which take into account agents' shared knowledge or public knowledge, such as common knowledge and generic common knowledge. These will be discussed in subsequent sections. The T axiom is a requirement of any system of epistemic logic. Without it, it would be possible to "know" false propositions, thus logics without factivity (T) such as $\mathsf{K4}_n$ ($\mathsf{S4}_n$ without T) may be seen to model belief but not knowledge. However, such systems, too, may be made more expressive by the addition of a common knowledge modality.

The following derived rule R3 is useful and also holds for the modalities that will be introduced: C for common knowledge and J for generic common knowledge.

Proposition 1. *R3 is a derived rule in any normal modal logic, i.e. one built on classical propositional logic with* K *and R2.*

$$\text{R3.} \quad \vdash \alpha \to \beta \ \Rightarrow \vdash K_i\alpha \to K_i\beta \ .$$

Proof.

1. $\alpha \to \beta$
2. $K_i(\alpha \to \beta)$ from 1. by R2

3. $K_i(\alpha \to \beta) \to (K_i\alpha \to K_i\beta)$ $\quad$ K
4. $K_i\alpha \to K_i\beta$ $\quad$ from 2. and 3. by R1.

$\square$

Definition 3. A *Kripke model* M for $\mathsf{S4}_n$ is $M = \langle W, R_1, R_2, \ldots, R_n, \Vdash \rangle$ where for $i \in \{1, 2, \ldots, n\}$

- W is a non-empty set of worlds;
- $R_i \subseteq W \times W$ is reflexive and transitive;
- $\Vdash \subseteq W \times \mathit{Var}$ so that for $w \in W$, $p \in \mathit{Var}$, $w \Vdash p$ iff p holds at w ;
- $\Vdash$ is extended to correspond with Boolean connectives at each world and so the accessibility relations correspond to the modalities: R_i corresponds to K_i so that in M,

$$u \Vdash K_i\varphi \text{ iff } (\forall v \in W)(uR_iv \Rightarrow v \Vdash \varphi) \ .$$

The accessibility relations R_i in T_n-models are reflexive while in $\mathsf{S5}_n$-models they are equivalence relations.

The worlds $\{v \mid wR_iv\}$ are those states agent i considers possible when at world w. This semantics of epistemically possible worlds is due to Hintikka ([Hin62]). While these logics are complete with respect to their respective models, it is of course also possible to have a model of a logic exist on a frame (= model/$\Vdash$) that is not “its own.”

Theorem 1. *T_n, $\mathsf{S4}_n$, and $\mathsf{S5}_n$ are sound and complete with respect to their models (see [BdRV01], or [Kri59] for the original single-agent proof).*

We show a standard completeness proof for $\mathsf{S4}_n$ as it is a basis for others we will present.

Proof of Soundness. Let M be an arbitrary $\mathsf{S4}_n$-model. We want to show $\mathsf{S4}_n \vdash \chi \Rightarrow M \Vdash \chi$. Assume χ is provable. To show χ holds in each world of M, it is enough to show that all the axioms and rules of $\mathsf{S4}_n$ are valid, by induction on χ.

- χ is a propositional variable: $u \Vdash \chi$ for all worlds in the model M implies χ is valid by definition.

- $\chi = \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi$. If χ is formed by Boolean connectives, it is valid by the definition of these connectives at each world.

- modus ponens: Suppose $u \Vdash \varphi \rightarrow \psi$. Then by the definition of the connectives, either $u \nVdash \varphi$ or $u \Vdash \psi$. If also $u \Vdash \varphi$, then $u \Vdash \psi$. So if $\varphi \rightarrow \psi$ and φ hold at any world, so does ψ.

- K axiom: $\chi = K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi) = (K_i(\varphi \rightarrow \psi) \wedge K_i\varphi) \rightarrow K_i\psi$. Suppose $u \Vdash K_i(\varphi \rightarrow \psi) \wedge K_i\varphi$, then for all v such that uR_iv,

$v \Vdash \varphi \to \psi$ and $v \Vdash \varphi$. So as modus ponens is valid $v \Vdash \psi$, and hence $u \Vdash K_i\psi$. Therefore $u \Vdash K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$ is valid.

- T axiom: $\chi = K_i\varphi \to \varphi$. Suppose $u \Vdash K_i\varphi$, then for all v such that uR_iv, $v \Vdash \varphi$. Since R_i is reflexive, uR_iu, and so $u \Vdash \varphi$. Thus $u \Vdash K_i\varphi \to \varphi$ is valid.

- $\mathsf{4}$ axiom: Suppose $u \Vdash K_i\varphi$, then for all v such that uR_iv, $v \Vdash \varphi$. As R_i is transitive, for all w such that vR_iw, uR_iw and so $w \Vdash \varphi$ and so $v \Vdash K_i\varphi$ and hence $u \Vdash K_iK_i\varphi$. Therefore $u \Vdash K_i\varphi \to K_iK_i\varphi$ is valid.

- modal necessitation: Assume φ is valid in M, then it is true at each world so $u \Vdash \varphi$, and for all worlds v such that vR_iu, $v \Vdash \varphi$. Thus $u \Vdash K_i\varphi$. As the world u was arbitrary, $K_i\varphi$ holds at all worlds and so is valid in the model. Therefore $\vdash \varphi \Rightarrow \vdash K_i\varphi$ is valid. □

Proof of Completeness. To show that $\mathsf{S4}_n$ is complete, we must show that each valid formula is provable, or the contrapositive, that each ϕ not provable in $\mathsf{S4}_n$ is not valid, i.e., fails to hold in some world of some $\mathsf{S4}_n$-model. A usual approach is to construct the "canonical model" which witnesses all provable formulas.

Definition 4. The *canonical model* for $\mathsf{S4}_n$ is $M' = \langle W, R_1, \ldots, R_n, \Vdash\rangle$ where

- $W = \{\Gamma \mid \Gamma$ is a maximally consistent set of $\mathsf{S4}_n$-formulas$\}$;

- $\Gamma R_i \Delta$ iff $\Gamma^i \subseteq \Delta$, where $\Gamma^i := \{\varphi \mid K_i\varphi \in \Gamma\}$;

- $\Vdash \subseteq W \times \mathit{Var}$ such that $\Gamma \Vdash p$ iff $p \in \Gamma$.

Proposition 2. *M' is indeed a model of $\mathsf{S4}_n$.*

Proof. As each Γ is maximally consistent and $\mathsf{S4}_n \vdash K_i\varphi \to \varphi$ (T), $\Gamma^i \subseteq \Gamma$ so $\Gamma R_i \Gamma$, thus R_i is reflexive. Suppose $\Gamma R_i \Delta$ and $\Delta R_i \Theta$. For each φ in Γ^i, $K_i\varphi \in \Gamma^i$ as well, since $\mathsf{S4}_n \vdash K_i\varphi \to K_iK_i\varphi$ ($\mathsf{4}$) and Γ is maximally consistent. Thus as $\Gamma R_i \Delta$, not only $\Gamma^i \subseteq \Delta$ but also $\Gamma^i \subseteq \Delta^i$, so as $\Delta R_i \Theta$, we have $\Gamma R_i \Theta$ as well, showing that R_i is transitive. $\square$

Lemma 1 (Truth Lemma)**.** *M' satisfies the Truth Lemma: for all Γ*

$$M', \Gamma \Vdash \varphi \Leftrightarrow \varphi \in \Gamma\ .$$

Proof. By induction on φ.

- base case: $\varphi = p$ for $p \in \mathit{Var}$. Holds by definition of $\Vdash$.

- Boolean cases: by extension of $\Vdash$, the induction hypothesis, and maximality of Γ.

- modal case: $\varphi = K_i\varphi$ ($\Leftarrow$) Assume $K_i\varphi \in \Gamma$. Then for all Δ such that $\Gamma R_i \Delta$, $\varphi \in \Delta$ so by the induction hypothesis, $\Delta \Vdash \varphi$. Thus $K_i\varphi \Vdash \Gamma$.

> ($\Rightarrow$) Assume $K_i\varphi \notin \Gamma$. Then $\Gamma^i \cup \{\neg\varphi\}$ must be consistent by the maximality of Γ, for otherwise φ would be provable and hence (by necessitation) so would $K_i\varphi$, which would contradict the consistency of Γ. If Δ is any maximally consistent set containing $\Gamma^i \cup \{\neg\varphi\}$, then $\Gamma R_i \Delta$ by definition of R_i. So $\Gamma \not\Vdash K_i\varphi$. $\square$

To finish completeness, assume $\mathsf{S4}_n \not\vdash \phi$. As $\{\neg\phi\}$ has some maximal consistent extension Φ, by the Truth Lemma, $\Phi \Vdash \neg\phi$ and so $\Phi \not\Vdash \phi$. $\square$

1.2 Common Knowledge and Generic Common Knowledge

Common knowledge is one standard tool for analyzing epistemic scenarios to determine what pieces of knowledge may be concluded by each agent in a particular situation. Three typical such scenarios are those of Muddy Children, Wise Men, and Unfaithful Wives. A history of these in particular is presented in opening chapter of [FHMV95].

In Muddy Children, a (finite) group of kids is playing outside and some of them get dirty. Everyone can see everyone else but they do not know whether they themselves are dirty. A parent says to them all, “At least one of you has a muddy forehead.” This statement is interesting because (except in the case of only one dirty kid) despite saying something they already know, this fact is now common knowledge, which it was not before. The parent then asks, "Do you know whether or not you are dirty?" They each say "yes" or "no" in unison and also hear everyone's answers. The parent asks the question again, and the children reply again, repeatedly, until eventually each child knows whether or not they themselves are dirty. The puzzle in this case is to determine how this happens in each set-up with $k \geq 1$ kids, $m \geq 1$ of whom are muddy. Of course, we assume that these children are proficient at epistemic reasoning and have a long attention span. We analyze a minimal form of this puzzle with two children in §2.2.1.

All of these classic puzzles involve nested knowledge operators, e.g., x knows that (y knows that (z knows that (x knows ψ))) and typically employ common knowledge of some fact to start the chain of reasoning.

The name generic common knowledge was suggested by Artemov to capture a state of a multi-agent epistemic system that yields iterated knowledge $I(\varphi)$: 'any agent knows that any agent knows that any agent knows… φ' for any number of iterations. The generic common knowledge of φ, $G.C.K.(\varphi)$, yields $I(\varphi)$,

$$G.C.K.(\varphi) \to I(\varphi)$$

but is not necessarily logically equivalent to $I(\varphi)$. Modal logics with $G.C.K.$

were suggested by McCarthy and Artemov and will be analyzed in the next chapter. These are in contrast to the prevalent formulation of Common Knowledge C due to Aumann in [Aum76] for which

$$C(\varphi) \leftrightarrow I(\varphi) \ .$$

The 'if' direction is secured for C as it is subject to the induction rule

$$\vdash \varphi \rightarrow E(\psi \wedge \varphi) \ \Rightarrow \vdash \varphi \rightarrow C\psi$$

([Bar88]), where $E\varphi$ is 'everybody knows φ.' It has been shown that in the usual epistemic scenarios cited, *G.C.K.* can replace the conventional common knowledge C ([Ant07]). Artemov noticed that such epistemic actions as public announcements of atomic sentences, generally speaking, yield *G.C.K.* rather than the conventional common knowledge ([Art12b]) and argues in [Art10] that in the analysis of perfect information games in the belief revision setting, Aumann's "no irrationality in the system" condition is fairly represented by a generic common knowledge rather than conventional common knowledge, and that this distinction lies at the heart of the well-known Aumann–Stalnaker controversy. We assume that the aforementioned arguments provide sufficient motivation for mathematical logical studies of the generic common knowledge and its different forms.

In this work, in addition to comparing these two forms of common knowledge in the implicit setting, we will describe a logic with explicit $G.C.K.$ and show that it has a corresponding modal system, i.e., the $G.C.K.$ modality along with the individual knowledge modalities, can be always made explicit.

1.3 Implicit and Explicit Knowledge

In modal logics, propositions and sentences may be known, but their reasons are not and so they model implicit knowledge. By contrast, in justification logics, we have formulas $t{:}\varphi$ where propositions and sentences φ are known explicitly by their proofs encoded in the proof term t. The oldest justification logic LP, logic of proofs, was developed with arithmetic provability semantics ([Art01], with preliminary work in 1995). Modals have been studied since Aristotle; the formalizations and axioms we will be using are from the early 20th century ([Bal10]). There is a close relationship between these two approaches: several modal logics can be said to have a direct explicit counterpart. If, in a justification logic all proof term are replaced by K, modal formulas result. If this is done for LP, the results are all $\mathsf{S4}$-compliant. A deeper result is that the other direction, *realization*, is also possible ([Art01]): by examining a modal derivation, proof terms can be recovered from modalities in such a way that a result is a theorem of a particular justification logic.

By Realizing S4 in LP, Artemov provided the exact provability semantics for S4, and hence for intuitionistic propositional logic, that Gödel had sought ([Art01]).

Chapter 2

C and *G.C.K.* in Modal Logic

Here we consider the relative strengths of three formal approaches to public knowledge: any fool knowledge by McCarthy et alia [McSHI78], common knowledge as presented by Halpern and Moses [HM90], and justified knowledge by Artemov [Art04]. We show that epistemic systems with the common knowledge modality C are conservative with respect to generic common knowledge systems with modality J on sentences of the form $\chi \wedge C\varphi \rightarrow \psi$, where χ, φ, and ψ are C-free, i.e. those which assume but do not conclude common knowledge formulas.

The formalization of common knowledge reflected in [HM90] traces back to Aumann in [Aum76] and has become standard in game-theoretic literature. Lewis's presentation of common knowledge in [Lew69], though informal, is reflected by McCarthy et alia and Artemov. These latter two turn out to be equivalent, though they arose from different developments and motivations.

This is the form we will refer to as generic common knowledge.

Look to §1.1 for multi-agent epistemic systems that will serve as the basis for the various common knowledge systems now presented. We will compare their logical strengths, semantics, and complexity, and will see that justified knowledge (J) systems are sufficient to solve classical epistemic scenarios, a role usually designated for common knowledge (C).

2.1 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 [Lew69] on convention. The intuition behind the informal definition of common knowledge below derives from Aumann's oft-cited [Aum76], where it was used in the context of agents having common priors. McCarthy's 'any fool' operator of 1970 ([FHMV95], p. 13) is closely related to common knowledge and his systems in [McSHI78] may have been 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 [HM90] (an expansion of a 1984 work of the same title) and Lehmann [Leh84]. Much of the work by Halpern and Moses appears in [FHMV95]. Common knowledge continues to be actively investigated.

Informally, the epistemic operator $C\varphi$, to be read 'φ is common knowledge,' is equivalent to the infinite conjunction iterated knowledge $I(\varphi)$:

$$C\varphi \leftrightarrow I(\varphi) = \varphi \wedge E\varphi \wedge EE\varphi \wedge E^3\varphi \wedge \cdots \wedge E^m\varphi \wedge \cdots$$

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, $\mathsf{S4}$, $\mathsf{S5}$, or other logic as appropriate. One formal characterization which [FHMV95] and [vBS04] take is via the Fixed Point Axiom

$$C\varphi \leftrightarrow E\,(\varphi \wedge C\varphi) \tag{2.1}$$

and the Induction Rule

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

yielding $C\varphi$ as the greatest fixed point solution to $X \leftrightarrow E(\varphi \wedge X)$ [FHMV95, vBS04]. 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 [Bar88]. [Gea92] too, provides a survey with examples but does not include a distinct formalism. There is also an equivalent axiomatic formulation of common knowledge in

[MvdH95] which replaces the induction rule with the induction axiom which, for technical convenience, we will use.

Definition 5. The *axioms and rules of* $\mathsf{S4}_n^C$:

CLASSICAL PROPOSITIONAL LOGIC:

A. axioms of classical propositional logic

R1. modus ponens

$\mathsf{S4}$ AXIOMS FOR ALL AGENTS $K_i,\ i \in \{1, 2, \ldots n\}$:

K. $K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$

T. $K_i\varphi \to \varphi$

$\mathsf{4}$. $K_i\varphi \to K_iK_i\varphi$

AXIOMS FOR COMMON KNOWLEDGE C:

K. $C(\varphi \to \psi) \to (C\varphi \to C\psi)$

T. $C\varphi \to \varphi$

C. $C\varphi \to E(C\varphi)$, where $E\varphi = \bigwedge_{i=1}^{n} K_i\varphi$ is read *everybody knows* φ

IA. $\varphi \wedge C(\varphi \to E\varphi) \to C\varphi$ *induction axiom*

NECESSITATION FOR ALL MODALS $\Box \in \{C, K_i\}$:

R2. $\vdash \varphi \Rightarrow \vdash \Box\varphi$.

For T_n^C or $\mathsf{S5}_n^C$, the K_i are all T or $\mathsf{S5}$ modalities, respectively.

Definition 6. A model M^C for $\mathsf{S4}_n^C$ is $M^C = \langle W, R_1, R_2, \ldots, R_n, R_C, \Vdash\rangle$ where

- $M = \langle W, R_1, R_2, \ldots, R_n, \Vdash\rangle$ is an $\mathsf{S4}_n$ model;

- $R_C = \big(\bigcup\limits_{i=1}^{n} R_i\big)^{\mathrm{TC}}$ is the transitive closure of all agents' relations;
- $\Vdash$ forcing relation is extended to all formulas so R_C corresponds to C:

$$u \Vdash C\varphi \text{ iff } (\forall v \in W)(uR_Cv \Rightarrow v \Vdash \varphi)\ .$$

R_C is reachability in the model; if you can get from w to u via a path of some R_i at each step, then wR_Cu. Models of T_n^C or $\mathsf{S5}_n^C$ are obtained when M is a model of T_n or $\mathsf{S5}_n$, respectively.

Theorem 2. T_n^C*,* $\mathsf{S4}_n^C$*, and* $\mathsf{S5}_n^C$ *are sound and complete with respect to their models (cf. [FHMV95], p. 70ff, [MvdH95], 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 [FHMV95], p. 93.

Recall the claim that common knowledge C is iterated knowledge I. This has even been considered definitional in limited situations, e.g. [FHMV95] p. 417. It is simple to see that $C\varphi$ implies each conjunct of iterated knowledge by using T for C, the following claim, and necessitation for all K_i.

Claim. $\mathsf{T}_n^C \vdash C\varphi \to E\varphi$.

Proof.

1. $C\varphi \to \varphi$ — T for C
2. $K_iC\varphi \to K_i\varphi$ — from 1. by R3 for K_i
3. $C\varphi \to K_iC\varphi$ — from C by definition of E
4. $C\varphi \to K_i\varphi$, for all i — from 3. and 2.
5. $C\varphi \to E\varphi$ — from 4. by definition of E.

This proof is sufficient for $\mathsf{S4}_n^C$ and $\mathsf{S5}_n^C$ as well. □

That $C\varphi$ implies $I\varphi$ can also be seen semantically by R_C. Assume $u \Vdash C\varphi$, then φ holds at all reachable worlds. This immediately secures $u \Vdash \varphi$ as R_C is reflexive, and $u \Vdash E\varphi$ as each $R_i \subseteq R_C$. As φ holds in all worlds v reachable from u, it holds in all worlds w reachable from u via v. So as each $R_i \subseteq R_C$, $v \Vdash E\varphi$ and hence $u \Vdash EE\varphi$. Considering that φ holds in all world reachable in m steps guarantees $u \Vdash E^m\varphi$.

To show that $I\varphi$ implies $C\varphi$ we can again take two approaches. Semantically, this again arises from the definition of R_C and syntactically, from consideration of the induction axiom (or rule). Semantically, suppose that iterated knowledge holds at u so that $u \Vdash E^k\varphi$ holds for all $k \geq 0$. Then $E^{k-1}\varphi$ holds at all worlds v accessible from u, for all R_i, and $E^{k-2}\varphi$ must hold at all worlds reachable from those v in one step by all R_i, or from u in two steps by all combinations of any R_i followed by any R_j. Extending this idea, we can see that by following all paths of length l from world u we find ourselves in a world w at which $E^{k-l}\varphi$ holds. As each R_i is reflexive, l is unbounded, as is k unbounded. Thus these worlds w are precisely the cluster of worlds reachable from u, i.e. $\{w \mid uRw,\ R = \big(\bigcup_{i=1}^{n} R_i\big)^{\mathrm{TC}} = R_C\}$. As $u \Vdash E^k\varphi$, $w \Vdash \varphi$ for all w such that $uRw = uR_Cw$, thus by completeness $u \Vdash C\varphi$.

2.2 Generic Common Knowledge

The formalism of generic common knowledge we first present was originally called justified knowledge. Justified knowledge was introduced by Artemov in [Art04, Art06] as the 'forgetful projection' of evidence-based knowledge represented by an adaptation of LP (Logic of Proofs). LP was the first in the now flourishing area of justification logic. In hybrid modal-LP systems (such as $\mathsf{T}_n\mathsf{LP}$, $\mathsf{S4}_n\mathsf{LP}$, $\mathsf{S5}_n\mathsf{LP}$) each formula may have both knowledge modalities and proof terms, such as in $\gamma \wedge K_2(\varphi \vee s : \psi) \rightarrow t : (K_1\chi)$. Modal justified knowledge systems (such as T^J_n, $\mathsf{S4}^J_n$, $\mathsf{S5}^J_n$) are ones in which all proofs are 'forgotten' and identified as one, so that the previous formula becomes $\gamma \wedge K_2(\varphi \vee J\psi) \rightarrow J(K_1\chi)$. Whereas $C\varphi$ asserts that φ is common knowledge, $J\varphi$ asserts that φ is common knowledge arising from a proof of φ or 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 due to the important Realization Theorem which provides an algorithm to reconstruct LP proof terms in $\mathsf{S4}_n\mathsf{LP}$ from $\mathsf{S4}_n^J$ derivations (and likewise for T_n and $\mathsf{S5}_n$) ([Art06]). Chapter 3 builds on this result with an algorithm to realize all K_i modalities, not just J, into an LP setting.

As with the logics with C, the construction of the generic common knowledge systems 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 may be chosen independently to be weaker, stronger, or the same as that of the agents'. In the aforementioned logics, J will be chosen to be an $\mathsf{S4}$ modality unless otherwise specified.

Definition 7. The *axioms and rules of* $\mathsf{S4}_n^J$:

CLASSICAL PROPOSITIONAL LOGIC:

A. axioms of classical propositional logic

R1. modus ponens

$\mathsf{S4}$ PRINCIPLES FOR ALL $K_i,\ i \in \{0, 1, 2, \ldots n\}$:

(J may be used in place of K_0):

K. $K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$

T. $K_i\varphi \to \varphi$

4. $K_i\varphi \to K_iK_i\varphi$

R2. $\vdash \varphi \Rightarrow \vdash K_i\varphi$

CONNECTION PRINCIPLE:

Con. $J\varphi \to K_i\varphi$.

For T^J_n or $\mathsf{S5}^J_n$, the K_i are all T or $\mathsf{S5}$ modalities, respectively, while J remains an $\mathsf{S4}$ modality.

At this point we can see a distinction between C and J in that the axioms for J may be chosen independently from the logical strength of the K_i.

Definition 8. A *model* M^J for $\mathsf{S4}^J_n$ is $M^J = \langle W, R_1, R_2, \dots, R_n, R_J, \Vdash\rangle$ where

- $M = \langle W, R_1, R_2, \dots, R_n, \Vdash\rangle$ is an $\mathsf{S4}_n$ model;
- $R_J \subseteq W \times W$ is reflective and transitive such that $R_J \supseteq \big(\bigcup_{i=1}^{n} R_i\big)^{\mathrm{TC}}$;
- $\Vdash$ forcing relation is extended to all formulas so R_J corresponds to J:

$$u \Vdash J\varphi \text{ iff } (\forall v \in W)(uR_Jv \Rightarrow v \Vdash \varphi) \ .$$

Models of T^J_n or $\mathsf{S5}^J_n$ are obtained when M is a model of T_n or $\mathsf{S5}_n$, respectively.

To distinguish R_J from R_C, consider a model M of $\mathsf{S4}_n$ with two disconnected components. Adding a relation R_C to M would keep it disconnected but adding R_J to M allows for connections between worlds in different components. If R_J is the total relation, we refer to that generic common knowledge as universal knowledge. If C is the greatest fixed point of $X \leftrightarrow E(\varphi\wedge X)$ ([FHMV95]), and J is any solution (see Proposition 6), then universal knowledge can be seen as the least fixed point.

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 [Art06].*

Proof of Soundness. Soundness for $\mathsf{S4}_n^J$ follows easily from the proof of Theorem 1. It only remains to check the connection principle.

- Con axiom: $\chi = J\varphi \to K_i\varphi$. Suppose $u \Vdash J\varphi$ so that for all v such that uR_Jv, $v \Vdash \varphi$. For all i, $R_i \subseteq R_J$ by definition, so for all w such that uR_iw, also uR_Jw and so $w \Vdash \varphi$, thus $u \Vdash K_i\varphi$. □

Note that to fully illustrate soundness of $\mathsf{S5}_n^J$, the soundness of axiom $\mathsf{5}$ would also need to be shown. This is straightforward when recalling that the R_is in $\mathsf{S5}_n^J$ are symmetric. For completeness of $\mathsf{S5}_n^J$, Artemov in [Art06] demonstrates a modal approach while for T_n^J and $\mathsf{S4}_n^J$, completeness is shown via equivalence to an appropriate Gentzen sequent system, which is needed for other purposes. For thoroughness, we here present the modal approach to the completeness proof of $\mathsf{S4}_n^J$.

Proof of Completeness. We demonstrate $\mathsf{S4}_n^J$-completeness via its canonical model.

Definition 9. The *canonical model* for $\mathsf{S4}_n^J$ is $M^{J\prime} = \langle W, R_1, \ldots, R_n, R_J \Vdash\rangle$ where for $i \in \{1, 2, \ldots, n\}$

- $W = \{\Gamma \mid \Gamma$ is a maximally consistent set of $\mathsf{S4}_n$-formulas$\}$;
- $\Gamma R_i \Delta$ iff $\Gamma^i \subseteq \Delta$, where $\Gamma^i := \{\varphi \mid K_i\varphi \in \Gamma\}$;
- $\Gamma R_J \Delta$ iff $\Gamma^J \subseteq \Delta$, where $\Gamma^J := \{\varphi \mid J\varphi \in \Gamma\}$;
- $\Vdash \subseteq W \times \mathit{Var}$ such that $\Gamma \Vdash p$ iff $p \in \Gamma$ for $p \in \mathit{Var}$.

Proposition 3. *M^J is indeed a model of $\mathsf{S4}_n^J$.*

Proof. As in the $\mathsf{S4}_n$ case, the R_i, and now R_J, are transitive and reflexive. It remains to be checked that $R_i \subseteq R_J$. Note that $\mathsf{S4}_n^J \vdash J\varphi \to K_iJ\varphi$ as a direct consequence of the $\mathsf{4}$ and Con axioms. Suppose $\Gamma R_i \Delta$ and $J\varphi \in \Gamma$. Then as Γ is maximally consistent $K_iJ\varphi \in \Gamma$ and so $J\varphi \in \Delta$. By the T axiom and the maximal consistency of Δ, $\varphi \in \Delta$. Thus $\Gamma R_J \Delta$. □

The Truth Lemma for M^J states $M^J, \Gamma \Vdash \varphi$ iff $\varphi \in \Gamma$ for any $\mathsf{S4}_n^J$-formula φ. As before, the proof is by induction on the complexity of φ with the additional modal case for J which is analogous to that of K_i.

To finish completeness, assume $\mathsf{S4}_n^J \nvdash \phi$. As $\{\neg\phi\}$ has some maximal extension Φ, by the Truth Lemma, $\Phi \Vdash \neg\phi$ and so $\Phi \nVdash \phi$. □

Recall that in common knowledge models, $R_C = \big(\bigcup_{i=1}^{n} R_i\big)^{\mathrm{TC}}$ 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 would seem that, if φ contains no Js, $J\varphi \to C\varphi$

but not the converse. This will be Proposition 8 at the end of the chapter but in this context we now we have the following proposition.

Definition 10. Let $\varphi^\dagger$ be φ with each instance of a J replaced by a C.

Proposition 4. $(\mathsf{S4}_n^J)^\dagger \subset \mathsf{S4}_n^C$ *but* $(\mathsf{S4}_n^J)^\dagger \neq \mathsf{S4}_n^C$.

Proof. It needs to be shown that the †-translation of each each rule and axiom of $\mathsf{S4}_n^J$ is provable in $\mathsf{S4}_n^C$, see Section 7 of [Art06] which uses the equivalent axiomatization of $\mathsf{S4}_n^C$ from [FHMV95]. It is only the induction axiom of $\mathsf{S4}_n^C$ which is not provable in $(\mathsf{S4}_n^J)^\dagger$, yielding strict inclusion. $\square$

Indeed, from Proposition 4, $(\mathsf{S4}_n^J)^\dagger \vdash (J\varphi)^\dagger \Rightarrow \mathsf{S4}_n^C \vdash C\varphi$, but not the converse.

Recall from the discussion at the end of §2.1 that $C\varphi$ implies iterated knowledge $I\varphi$. The same arguments hold to show $J\varphi$ implies $I\varphi$, or use the above proposition showing that $(J\varphi)^\dagger$ implies $C\varphi$. Iterated knowledge does not, however, imply generic common knowledge as $I\varphi \leftrightarrow C\varphi \not\rightarrow (J\varphi)^\dagger$.

An analog of Proposition 4 for T and $\mathsf{S5}$ holds as well. An $\mathsf{S5}$ justified knowledge logic where J is an $\mathsf{S5}$ modality, i.e. $\mathsf{S5}_n^{J(\mathsf{S5})}$, is considered in [Rub06] 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* IA *be the induction axiom* $\varphi \wedge C(\varphi \to E\varphi) \to C\varphi$*. Then*

$$\mathsf{S4}_n^C \equiv (\mathsf{S4}_n^J)^\dagger + \mathsf{IA}\ , \quad \mathsf{T}_n^C \equiv (\mathsf{T}_n^J)^\dagger + \mathsf{IA}\ , \text{ and } \quad \mathsf{S5}_n^C \equiv (\mathsf{S5}_n^{J(\mathsf{S5})})^\dagger + \mathsf{IA}\ .$$

Proof. The strict inclusion of the J systems follows from Proposition 4 and the observation 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. □

2.2.1 Muddy Children example

We now work through an instance of the Muddy Children scenario in $\mathsf{S4}_2^J$ in which there are two children, called a and b, both of whom are muddy. Assume at the start neither knows they are muddy though they can see that the other is. A parent sees them and says, "At least one of you is muddy. Do you know whether or not you are muddy?" to which they both answer, "No, I don't know." The derivation below shows how after hearing these statements, child a is able to conclude that she is muddy. Let A be the proposition that a is muddy and B be the proposition that b is muddy.

The context $\mathcal{S}$ for child a after they both declare, "I don't know," is:

$$\mathcal{S} = J(A \vee B) \wedge J(K_b \neg A \vee K_b A) \wedge J(\neg K_b B) \wedge J(\neg K_b \neg B)\ .$$

In derivations from $\mathcal{X}$, $\mathsf{S4}_2^J$, the only postulated rule is modus ponens

(R1), while Necessitation (R2) is, generally speaking, not admissible. Otherwise, epistemic soundness can be violated, e.g., by assuming $\mathcal{X}$ we mean that $\mathcal{X}$ is true, which does not justify concluding that $\mathcal{X}$ is known, $J\mathcal{X}$.

However, in this case, since all conjuncts in $\mathcal{S}$ are modalized by J, for derivations in $\mathcal{S}$, $\mathsf{S4}_2^J$ the rule of Necessitation for all the modalities $\Box \in \{K_1, K_2, J\}$ is admissible and we will use them freely. This can be justified by a straightforward induction on the derivations in $\mathcal{S}$, $\mathsf{S4}_2^J$. Indeed, if X is an axiom of $\mathsf{S4}_2^J$, then $\Box X$ is derivable in $\mathcal{S}$, $\mathsf{S4}_2^J$. If X is $\mathcal{S}$, then $\mathsf{S4}_2^J$ derives $X \to \Box X$ (here we rely on the specific structure of this $\mathcal{S}$), hence given X we can conclude $\Box X$ by just modus ponens. Finally, if X is obtained from Y and $Y \to X$ by modus ponens, by the I.H., $\Box Y$ and $\Box(Y \to X)$ are also given, hence by $\mathsf{S4}$ reasoning, $\Box X$.

Claim. *Given $\mathcal{S}$, child a can conclude she is muddy: $\mathcal{S}, \mathsf{S4}_2^J \vdash K_a A$.*

Proof.

1. $(A \vee B) \to (\neg A \to B)$ propositional tautology
2. $J(A \vee B) \to J(\neg A \to B)$ from 1. by R3 for J
3. $J(A \vee B)$ from $\mathcal{S}$
4. $J(\neg A \to B)$ from 2. and 3. by R1
5. $K_b(\neg A \to B)$ from 4. by Con and R1
6. $K_b \neg A \to K_b B$ from 5. and K for K_b by R1
7. $\neg K_b B \to \neg K_b \neg A$ from 6. by contraposition

8. $K_a(\neg K_b B) \to K_a(\neg K_b \neg A)$ from 7. by R3 for K_a
9. $J(\neg K_b B)$ from $\mathcal{S}$
10. $K_a(\neg K_b B)$ from 9. by Con and R1
11. $K_a(\neg K_b \neg A)$ from 8. and 10. by R1
12. $(K_b \neg A \vee K_b A) \to (\neg K_b \neg A \to K_b A)$ propositional tautology
13. $J(K_b \neg A \vee K_b A) \to J(\neg K_b \neg A \to K_b A)$ from 12. by R3 for J
14. $J(K_b \neg A \vee K_b A)$ from $\mathcal{S}$
15. $J(\neg K_b \neg A \to K_b A)$ from 13. and 14. by R1
16. $K_a(\neg K_b \neg A \to K_b A)$ from 15. by Con and R1
17. $K_a(\neg K_b \neg A) \to K_a(K_b A)$ from 16. and K for K_a and R1
18. $K_a(K_b A)$ from 11. and 17. by R1
19. $K_a(K_b A) \to K_a A$ by R3 on T axiom for K_b
20. $K_a A$ from 19. and 20. by R1 .

□

From Corollary 1, a similar derivation could be made in $\mathsf{S4}_n^C$ (changing Js to Cs), and in fact this derivation only made use of T_2^J. Of course there are assumptions made when modeling these classic puzzles such as a's ability to reason about what b knows. The parent's statement to a and b together is a public announcement of a fact. Such public announcements are the standard method of obtaining common knowledge. Here we relax this to generic common knowledge as it is sufficient for the derivation. At the end of the chapter, we will see conservativity results which state that J is sufficient whenever we just want to use, and not conclude, C. In [Art12b], Artemov points out that public announcements, far from producing common knowledge, actually yield universal knowledge.

The evidence-based common knowledge semantics for J systems are enriched by the Realization Theorem mentioned at the start of the section. This gives a constructive approach to recovering, Realizing, the full proof terms for J in hybrid implicit/explicit evidence-based knowledge systems.

Theorem 4 (Realization Theorem)**.** *(Artemov) 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*.*

Both Theorem 4 and a theorem for realizing $\mathsf{S5}_n^J$ in $\mathsf{S5}_n\mathsf{LP}$ are established in [Art06]. A Realization Theorem for $\mathsf{S5}_n^{J(\mathsf{S5})}$ is given in [Rub06]. The first such Realization Theorem appears in [Art01] and shows LP to be the exact explicit analog of $\mathsf{S4}$. Theorem 12 builds on Theorem 4, realizing each modality K_i, as well as J, in a new logic $\mathsf{LP}_n(\mathsf{LP})$.

In addition to the semantic and axiomatic distinction between J and C, there is the issue of complexity. In systems with J, the cut rule can be eliminated ([Art06]) but not for those with C ([AJ05]). The complexity of $\mathsf{S4}_n$ for $n \geq 1$ is *PSPACE-complete* and $\mathsf{S4}_n^C$ for $n \geq 2$ is *EXPTIME*-complete ([FHMV95]). The complexity $\mathsf{S4}_1^J$ is *PSPACE*-complete [Dem00], but some experts conjecture that the complexity for $n \geq 2$ might be lower than *EXPTIME*. These features are exploited by Bryukhov in [Bry05] to develop an automated theorem prover for $\mathsf{S4}_n^J$.

2.3 Any Fool's Knowledge

McCarthy's modeling of common knowledge via "any fool knows" apparently traces back to 1970 ([FHMV95], p. 13), though its first published appearance is in [McSHI78]. 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 [McSHI78] 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 alia use subscripted modals, S_i, $i \in \{0, 1, \ldots, n\}$, 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 11. The McCarthy et alia systems are built from:

CLASSICAL PROPOSITIONAL LOGIC:

A. axioms of classical propositional logic

R1. modus ponens

KNOWLEDGE AXIOMS:

link. $O\varphi \to S\varphi$ *what any fool knows, anyone knows*

K0. $S\varphi \to \varphi$ *if someone knows it, it is true*

K1. $O(S\varphi \to \varphi)$ *any fool knows that only true propositions can be known*

K2. $O(O\varphi \to OS\varphi)$ *any fool knows what he knows he knows anyone knows*

K3. $O(S\varphi \wedge S(\varphi \to \psi) \to S\psi)$ *any fool knows anyone can do modus ponens*

K4. $O(S\varphi \to SS\varphi)$ *any fool knows anyone can do positive introspection*

K5. $O(\neg S\varphi \to S\neg S\varphi)$ *any fool knows anyone can do negative introspection.*

We will look at three systems identified in [McSHI78] given by the link axiom and K0-K3, K0-K4, or K0-K5.[1] These will be referred to as MT, M4, and M5 respectively. Model semantics and completeness results for a variant of M5 is stated in [McSHI78]. 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 epistemic scenarios, of which Wise Men and Unfaithful Wives are addressed in [McSHI78]. These particular axioms do not seem to be built on standard formulations of modal logics

---

[1] In [McSHI78], K0 is omitted from these lists. Given other statements in the paper, this clearly is just an oversight.

but engineered for these scenarios, and 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 Theorem 5.

Definition 12. Let $\varphi^{\ddagger}$ be φ with each instance of a J replaced by an O and each K_i replaced by S_i.

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

Proof. Immediate from Lemmas 2–4 which follow. □

Lemma 2. $(\mathsf{T}_n^J)^{\ddagger} \equiv \mathsf{MT}$ *.*

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

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

link: $\mathsf{T}_n^J \vdash J\varphi \to K_i\varphi$; Con

K0: $\mathsf{T}_n^J \vdash K_i\varphi \to \varphi$; T for K_i

K1: $\mathsf{T}_n^J \vdash J(K_i\varphi \to \varphi)$; J necessitation of T of K_i

K2: $\mathsf{T}_n^J \vdash J(J\varphi \to JK_i\varphi)$

$\mathsf{T}_n^J \vdash J\varphi \to JJ\varphi$ 4 for J

$\mathsf{T}_n^J \vdash J\varphi \to K_i\varphi$ Con

$\mathsf{T}_n^J \vdash JJ\varphi \to JK_i\varphi$ from 2. by R3

$\mathsf{T}_n^J \vdash J\varphi \to JK_i\varphi$ from 1. and 3.

$\mathsf{T}_n^J \vdash J(J\varphi \to JK_i\varphi)$; from 4. by J necessitation

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

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

link′: $\mathsf{T}_n^J \vdash J\varphi \to J\varphi$; 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 for J

K2′: $\mathsf{T}_n^J \vdash J(J\varphi \to JJ\varphi)$; J necessitation of 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.

(⇒) $(\mathsf{T}_n^J)^\ddagger \subset \mathsf{MT}$. We must show that MT satisfies the $(\mathsf{T}_n^J)^\ddagger$ 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, link, 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 for S, O is a an S

4: $\mathsf{MT} \vdash O\varphi \to OO\varphi$; by K2, T for O, O is an S

Con: $\mathsf{MT} \vdash O\varphi \to S\varphi$; link.

O necessitation: This follows as each S and O axiom is necessitated.

K axiom for S and O is necessitated by K3.

T axiom for S and O is necessitated by K1.

4 axiom for O is necessitated by K2.

S necessitation: This follows from O necessitation and link. □

Lemma 3. $(\mathsf{S4}_n^J)^\ddagger \equiv \mathsf{M4}$.

Proof. Recall that J and K_i are $\mathsf{S4}$ modalities.

($\Leftarrow$) $(\mathsf{S4}_n^J)^\ddagger \supset \mathsf{M4}$ follows from Lemma 2 and

$\mathsf{K4}$: $\mathsf{S4}_n^J \vdash J(K_i\varphi \to K_iK_i\varphi)$; by J necessitation of $\mathsf{4}$ for K_i

$\mathsf{K4}'$: $\mathsf{K2}'$.

($\Rightarrow$) $(\mathsf{S4}_n^J)^\ddagger \subset \mathsf{M4}$ follows from Lemma 2 and

S axioms:

$\mathsf{4}$: $\mathsf{M4} \vdash S\varphi \to SS\varphi$; by $\mathsf{K4}$, T for O.

O necessitation: $\mathsf{4}$ for S is necessitated by $\mathsf{K4}$. □

Lemma 4. $(\mathsf{S5}_n^{J(\mathsf{S5})})^\ddagger \equiv \mathsf{M5}$.

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

($\Leftarrow$) $(\mathsf{S5}_n^{J(\mathsf{S5})})^\ddagger \supset \mathsf{M5}$ follows from Lemma 3 and

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

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

($\Rightarrow$) $(\mathsf{S5}_n^{J(\mathsf{S5})})^\ddagger \subset \mathsf{M5}$ follows from Lemma 3 and

S axioms:

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

O axioms:

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

O necessitation: 5 axiom for O and S necessitated by K5 and link. □

Lemma 4 completes the proof of Theorem 5. Despite quite different motivations, McCarthy's "any fool" and Artemov's justified knowledge approaches lead to the same multi-modal logics. It is interesting to compare the semantics for syntactically equivalent operators: whereas O is considered a fool, the lowest common denominator, J can be viewed by the other agents as a reliable authority, a trusted source of proofs.

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

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 for J and endows the O systems with a constructive, evidence-based semantics via the Realization Theorem and shows O, also, to be a generic common knowledge modality.

2.4 Limited Conservativity

A logic L with language $\mathcal{L}$ is a conservative extension of a logic L′ with language $\mathcal{L}' \subseteq \mathcal{L}$ if for sentences φ of $\mathcal{L}'$, L proves φ if and only if L′ proves φ. Recall † of Definition 10 which renames J to C. As the logics $(\mathsf{S4}_n^J)^\dagger$ and

$\mathsf{S4}_n^C$ have the same language and yet are not equal, it is clear that $\mathsf{S4}_n^C$ cannot be a conservative extension of $(\mathsf{S4}_n^J)^\dagger$, it is however a conservative extension over all formulas in which C occurs only negatively. Following the standard terminology, 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 an odd number of negation symbols. For example, X occurs only negatively in these first two formulas and both positively and negatively in the last:

$$X \to Y, \quad (\neg(A \wedge X) \to B) \to Y, \quad A \wedge X \to B \vee X \ .$$

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

In some sense this result is tight, as by Corollary 1 the induction axiom $(\varphi \wedge C(\varphi \to E\varphi) \to C\varphi)$ which distinguishes $(\mathsf{S4}_n^J)^\dagger$ from $\mathsf{S4}_n^C$ has, along with a negative occurrence of C, a single positive occurrence of C.

Proof. $(\Rightarrow)$ is secured by the inclusion $(\mathsf{S4}_n^J)^\dagger \subset \mathsf{S4}_n^C$ of Proposition 4.

$(\Leftarrow)$ This direction is a consequence of the Sign 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 $\mathsf{S4}_n^J \nvdash \varphi$. By completeness, there is a model M and a world x such that $M, x \Vdash \neg\varphi$. By the Sign Lemma, $M^\dagger, x \Vdash^\dagger \neg(\varphi)^\dagger$, hence $\mathsf{S4}_n^C \nvdash (\varphi)^\dagger$, since $M^\dagger$ (with R_J ignored) is a model for $\mathsf{S4}_n^C$. □

Lemma 5 (Sign 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 $M^\dagger$, where $\Vdash^\dagger$ 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 $x \Vdash \varphi \quad \Rightarrow \quad x \Vdash^\dagger (\varphi)^\dagger$;

if all occurrences of J in φ are negative, then $x \nVdash \varphi \quad \Rightarrow \quad x \nVdash^\dagger (\varphi)^\dagger$.

Proof. By induction on φ.

- base case: $\varphi = p$ is secured by the definition of $\Vdash^\dagger$.
- Boolean case: $\varphi = \psi \to \theta$.
  - All occurrences of J in φ are positive and $x \Vdash \varphi$. Then $x \nVdash \psi$ or $x \Vdash \theta$. In the former case all occurrences of J in ψ are negative and, by the induction hypothesis, $x \nVdash^\dagger (\psi)^\dagger$. In the latter case, all occurrences of J in θ are positive and, by the induction hypothesis, $x \Vdash^\dagger (\theta)^\dagger$. In either case, $x \Vdash^\dagger (\varphi)^\dagger$.
  - All occurrences of J in φ are negative and $x \nVdash \varphi$. Then $x \Vdash \psi$ and $x \nVdash \theta$. Since all occurrences of J in ψ are positive and all

occurrences of J in θ are negative, by the induction hypothesis, $x \Vdash^{\dagger} (\psi)^{\dagger}$ and $x \nVdash^{\dagger} (\theta)^{\dagger}$, hence $x \nVdash^{\dagger} (\varphi)^{\dagger}$.

- modal case: $\varphi = K_i\psi$.
  - 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, $y \Vdash^{\dagger} (\psi)^{\dagger}$ for all y such that xR_iy, hence $x \Vdash^{\dagger} (K_i\psi)^{\dagger}$, i.e., $x \Vdash^{\dagger} (\varphi)^{\dagger}$.
  - All occurrences of J in φ are negative and $x \nVdash \varphi$. Then for some y such that xR_iy, $y \nVdash \psi$. Since all occurrences of J in ψ are also negative, by the induction hypothesis, $y \nVdash^{\dagger} (\psi)^{\dagger}$, hence $x \nVdash^{\dagger} (K_i\psi)^{\dagger}$, i.e., $x \nVdash^{\dagger} (\varphi)^{\dagger}$.
- modal case: $\varphi = J\psi$.
  - 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, $y \Vdash^{\dagger} (\psi)^{\dagger}$ for all y such that xR_Cy. Hence $x \Vdash^{\dagger} C(\psi)^{\dagger}$, i.e., $x \Vdash^{\dagger} (J\psi)^{\dagger}$, i.e., $x \Vdash^{\dagger} (\varphi)^{\dagger}$.
  - ‘All occurrences of J in φ are negative and $x \nVdash \varphi$’ is impossible,

since $\varphi = J\psi$ and the displayed occurrence of J is positive in $J\psi$. $\square$

Corollary 3. *If χ, φ and ψ are formulas in the language of $\mathsf{S4}_n$, then*

$$\mathsf{S4}_n^C \vdash \chi \wedge C\varphi \rightarrow \psi \Leftrightarrow \mathsf{S4}_n^J \vdash \chi \wedge J\varphi \rightarrow \psi \ .$$

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

Corollary 4. *If χ, φ and ψ are formulas in the language of T_n, then*

$$\mathsf{T}_n^C \vdash \chi \wedge C\varphi \rightarrow \psi \Leftrightarrow \mathsf{T}_n^J \vdash \chi \wedge J\varphi \rightarrow \psi \ .$$

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

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 Sign Lemma starts with $\mathsf{S5}_n^{J(\mathsf{S5})}$-models ($J$ is an $\mathsf{S5}$ modality) and completeness for $\mathsf{S5}_n^{J(\mathsf{S5})}$. $\square$

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 $\varphi = J\psi$ requires only that $R_C \subseteq R_J$ and not that R_J be of the same logical strength. The proof actually applies to 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 occurances of C, $\mathsf{S5}_n^C$ is a conservative extension of $(\mathsf{S5}_n^J)^\dagger$.

This conservativity of C over J limited to formulas with C in negative position suggested the use of J in place of C in situations in which common knowledge is applied or assumed, rather than derived or concluded. In the Muddy Children example, the children's answers are justified by J after each round of questioning; we are not interested in concluding that anything is common knowledge for these children.

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. This we do now.

2.5 A joint logic of C and $G.C.K.$

A multi-agent epistemic logic $\mathsf{S4}_n^{CJ}$ is defined which weds $\mathsf{S4}_n^C$ and $\mathsf{S4}_n^J$ so that there are mixed formulas involving the common knowledge operator C and the generic common knowledge operator J. Completeness for this logic is shown, allowing for a direct comparison of the deductive strength of J and C and easing comparisons when considering typical applications of common knowledge such as in scenarios previously mentioned.

2.5.1 Axiomatization of $\mathsf{S4}_n^{CJ}$

We construct $\mathsf{S4}_n^{CJ}$ which encompasses the languages of and logics of $\mathsf{S4}_n^{C}$ and $\mathsf{S4}_n^{J}$. As it has J and C modalities, we can consider formulas which contain them both, and both of them with K_i modalities. The next subsection shows the soundness and completeness of $\mathsf{S4}_n^{CJ}$ with respect to $\mathsf{S4}_n^{CJ}$-models and some corollaries.

Definition 13. The *axioms and rules of* $\mathsf{S4}_n^{CJ}$, for $i \in \{1, 2, \ldots, n\}$:

CLASSICAL PROPOSITIONAL CALCULUS:

A. axioms of classical propositional calculus

R1. modus ponens

$\mathsf{S4}$ AXIOMS FOR ALL MODALITIES $K_i,\ J,\ C$

$\mathsf{S4}$ ADDITIONAL KNOWLEDGE AXIOMS:

Con. $J\varphi \rightarrow K_i\varphi$

ConC. $C\varphi \rightarrow K_i\varphi$

IA. $\varphi \wedge C(\varphi \rightarrow E\varphi) \rightarrow C\varphi$, where $E\varphi = \bigwedge_{i=1}^{n} K_n\varphi$

NECESSITATION FOR ALL MODALITIES, $\Box \in \{K_i,\ J,\ C\}$:

R2. $\vdash \varphi \Rightarrow\ \vdash \Box\varphi$.

Compare the following with Proposition 4.

Proposition 5. $\mathsf{S4}_n^{CJ} \vdash J\varphi \rightarrow C\varphi$.

Proof. Reasons from propositional calculus are not listed.

1. $J\varphi \to EJ\varphi$ from 4 and Con, definition of E
2. $C(J\varphi \to EJ\varphi)$ from 1. by R2 for C
3. $J\varphi \to C(J\varphi \to EJ\varphi)$ from 2.
4. $J\varphi \to J\varphi$
5. $J\varphi \to J\varphi \wedge C(J\varphi \to EJ\varphi)$ from 3. and 4.
6. $J\varphi \wedge C(J\varphi \to EJ\varphi) \to CJ\varphi$ IA on $J\varphi$
7. $J\varphi \to CJ\varphi$ from 5. and 6.
8. $J\varphi \to \varphi$ T for J
9. $CJ\varphi \to C\varphi$ from 8. by R3 for C
10. $J\varphi \to C\varphi$ from 7. and 9.

□

That the converse does not hold must wait till Proposition 8, after $\mathsf{LP}_n(\mathsf{LP})$ is shown to be sound and complete.

Proposition 6. *Both $C\varphi$ and $J\varphi$ satisfy X in (2.1), the Fixed Point Axiom,*

$$X \leftrightarrow E(\varphi \wedge X) \ .$$

Proof. $J\varphi \leftrightarrow E(\varphi \wedge J\varphi)$:

(→)

1. $JJ\varphi \to EJ\varphi$ from Con and definition of E
2. $J\varphi \to JJ\varphi$ 4 for J
3. $J\varphi \to EJ\varphi$ from 2. and 1.
4. $J\varphi \to E\varphi$ from Con and definition of E
5. $J\varphi \to (E\varphi \wedge EJ\varphi)$ from 3. and 4.
6. $J\varphi \to E(\varphi \wedge J\varphi)$ from 5., normal modalities[2] commute with $\wedge$

	1.	$E(\varphi \wedge J\varphi) \to E\varphi \wedge EJ\varphi$	normal modalities commute with $\wedge$
	2.	$E\varphi \wedge EJ\varphi \to EJ\varphi$	
$(\leftarrow)$	3.	$EJ\varphi \to K_iJ\varphi$	definition of E
	4.	$K_iJ\varphi \to J\varphi$	T for K_i
	5.	$E(\varphi \wedge J\varphi) \to J\varphi$	from 1. – 4.

As each J axiom or rule is also a C axiom or rule, since J satisfies the fixed point axiom, so does C. □

We will use the following proposition in the completeness proof.

Proposition 7. $\mathsf{S4}_n^{CJ} \vdash C\varphi \to EC\varphi$. *This is axiom* C *of* $\mathsf{S4}_n^{C}$.

Proof. Just as lines 1. – 3. in forward direction of proof of Proposition 6. □

2.5.2 Completeness of $\mathsf{S4}_n^{CJ}$

Definition 14. A *model* M^{CJ} *of* $\mathsf{S4}_n^{CJ}$ is $M^{CJ} = \langle W, R_1, \ldots, R_n, R_C, R_J, \Vdash \rangle$ such that

- $\langle W, R_1, \ldots, R_n, R_C, \Vdash \rangle$ is a model of $\mathsf{S4}_n^{C}$;
- $R_J \subseteq W \times W$ is reflexive and transitive and $R_C \subseteq R_J$;
- $\Vdash$ is extended so that R_J corresponds to J:

$$u \Vdash J\varphi \text{ iff } (\forall v \in W)(uR_Jv \Rightarrow v \Vdash \varphi) .$$

Theorem 7. $\mathsf{S4}_n^{CJ}$ *is sound and complete with respect to* M^{CJ} *models.*

Soundness. Let M be an arbitrary $\mathsf{S4}_n^{CJ}$-model. Assume χ is provable and show it holds in each world of M. It is enough to show that all the axioms and rules are valid. This follows from the proof of Theorem 3. It remains to check:

- ConC: Analogous to the proof shown for J following Theorem 3.
- IA: $\chi = \varphi \wedge C(\varphi \to E\varphi) \to C\varphi$. Suppose $u \Vdash \varphi \wedge C(\varphi \to E\varphi)$. Then for all v such that uR_Cv, $v \Vdash \varphi \to E\varphi$ $(**)$. We want to show $u \Vdash C\varphi$, i.e. $v \Vdash \varphi$ for all v reachable from u. Proceed by induction on length of path l along R_is from u to v. It is sufficient to show this for paths of length l along the R_is as then the R_C paths are of length $\leq l$ (and in fact of length 0 or 1 along R_C).

  – If $l = 0$ then $u = v$ and by assumption, $u \Vdash \varphi$.

  – induction hypothesis: Assume $s \Vdash \varphi$ holds for worlds s reachable from u by a path of length l.

  – Suppose that v is reachable from u by a path of length $l + 1$. Then there is a world t reachable from u in l steps and tR_iv for some i. By the induction hypothesis, $t \Vdash \varphi$ but also by $(**)$ and modus ponens, $t \Vdash E\varphi$. But tR_iv, so $v \Vdash \varphi$. Thus $u \Vdash C\varphi$. $\square$

To show completeness, the usual approach would be to construct the canonical model. However, here the canonical structure turns out not to be a model of $\mathsf{S4}_n^{CJ}$. So, instead of a single large model which acts as a counter-model for all non-provable ϕ, for each non-provable ϕ we construct a finite model with a world at which ϕ does not hold. We use the canonical structure and filtration techniques to construct these counter-models. The remainder of the proof of Theorem 7 is delayed until the end of this subsection after the presentation on filtrations.

Definition 15. The *canonical structure* for $\mathsf{S4}_n^{CJ}$ is $M^{CJ'} =$ $\langle W, R_1, \ldots, R_n, R_C, R_J, \Vdash \rangle$ where

- $W = \{\Gamma \mid \Gamma$ is a maximally consistent set of $\mathsf{S4}_n^{CJ}$ formulas$\}$;
- $\Vdash \subseteq W \times \mathit{Var}$ such that $\Gamma \Vdash p$ iff $p \in \Gamma$ for $p \in \mathit{Var}$;
- $\Gamma R_i \Delta$ iff $\Gamma^i \subseteq \Delta$, where $\Gamma^i := \{\varphi \mid K_i\varphi \in \Gamma\}$;
- $\Gamma R_J \Delta$ iff $\Gamma^J \subseteq \Delta$, where $\Gamma^J := \{\varphi \mid J\varphi \in \Gamma\}$;
- $\Gamma R_C \Delta$ iff $\Gamma^C \subseteq \Delta$, where $\Gamma^C := \{\varphi \mid C\varphi \in \Gamma\}$.

It follows from completeness proofs of $\mathsf{S4}_n^J$ and $\mathsf{S4}_n^C$ that $M^{CJ'}$ also satisfies the Truth Lemma: $M^{CJ'}, \Gamma \Vdash \varphi \Leftrightarrow \varphi \in \Gamma$.

Corollary 6. *As a consequence of the Truth Lemma, any maximal consistent set of formulas is satisfiable in* $M^{CJ'}$.

Thus $\mathsf{S4}_n^{CJ} \vdash \varphi \Rightarrow M^{CJ'}, \Gamma \Vdash \varphi$, so soundness holds for the canonical structure.

Lemma 6. *The canonical structure* $M^{CJ'}$ *is not a model of* $\mathsf{S4}_n^{CJ}$ *(cf. [MvdH95] p. 50).*

In $M^{CJ'}$, all accessibility relations are reflexive and transitive and $R_C \subseteq R_J$. However, $R_C \neq \big(\bigcup\limits_{i=1}^{n} R_i\big)^{\mathrm{TC}}$ as we only have $\big(\bigcup\limits_{i=1}^{n} R_i\big)^{\mathrm{TC}} \subset R_C$, thus $M^{CJ'}$ is not a model of $\mathsf{S4}_n^{CJ}$.

Proof. It suffices to show that $R_C \not\subset \big(\bigcup\limits_{i=1}^{n} R_i\big)^{\mathrm{TC}}$. Consider a set of formulas

$$\Phi = \{Ep, EEp, EEEp, \dots\} \cup \{\neg Cp\}$$

for some $p \in \mathit{Var}$ and abbreviate $EEEp$ as E^3p, etc.

Claim. Φ *is* $\mathsf{S4}_n^{CJ}$*-consistent.*

Proof of Claim. Suppose Φ is inconsistent. Then there is a finite $\Delta \subset \Phi$ which is already inconsistent so say $\Delta = \{E^{k_1}p, E^{k_2}p, \dots, E^{k_m}p \mid k_i < k_{i+1}\ for\ i < m\} \cup \{\neg Cp\}$. (If Δ were already inconsistent, including $\{\neg Cp\}$ would keep Δ inconsistent.) Consider the model $N = \langle W, R_1, R_2, R_C, R_J, \Vdash\rangle$ where

- $W = \mathbb{N}$;
- $R_1 = \{(n,n) \mid n \in \mathbb{N}\} \cup \{(n,n+1),(n+1,n) \mid n \in \mathbb{N} \text{ and } n \text{ even}\}$;
- $R_2 = \{(n,n) \mid n \in \mathbb{N}\} \cup \{(n,n+1),(n+1,n) \mid n \in \mathbb{N} \text{ and } n \text{ odd}\}$;
- $R_J = R_C = (R_1 \cup R_2)^{\mathrm{TC}}$;
- $x \Vdash p$ iff $x \leq k_m + 1$.

Figure 1 shows the frame of N with the reflexive arrows of R_1 and R_2 suppressed. For this model R_C is an equivalence relation with one class, mR_Cn for all $m,n \in \mathbb{N}$.

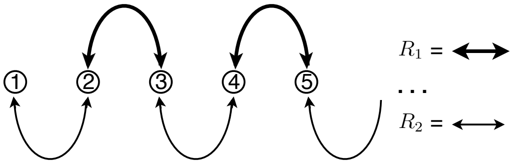

Figure 1.

But $N,1 \Vdash \Delta$. To see why, consider an example where $k_m = 3$ thus

$1,2,3,4 \Vdash p$,

$1,2,3 \Vdash K_1p \wedge K_2p \wedge Ep$ though $4 \Vdash \neg K_1p$,

$1,2 \Vdash K_1Ep \wedge K_2Ep \wedge EEp$ though $3 \Vdash \neg K_2K_1p$,

$1 \Vdash K_1EEp \wedge K_2EEp \wedge EEEp$ though $2 \Vdash \neg K_1K_2K_1p$, and

$1 \Vdash \neg C\varphi$ as $5 \Vdash \neg p$ and $1R_C5$.

Since $1 \Vdash E^3p \wedge \neg Cp$, this Δ is satisfied and hence is consistent. Since no finite subset Δ is inconsistent, Φ is consistent. $\square_{\mathit{Claim}}$

We now finish the proof of Lemma 6. Since Φ is consistent, it is contained in some maximal consistent set Φ'. Let $\Theta = \{\neg p\} \cup \{\theta \mid C\theta \in \Phi'\}$. Note that Θ is consistent. As $\{\theta \mid C\theta \in \Phi'\} \subseteq \Phi'$ which is maximal consistent, Θ could only be inconsistent if $\neg p \wedge p \in \Theta$. As $\neg Cp \in \Phi$, Cp is not in Φ', so p is not in Θ, so Θ is consistent, and so contained in some maximal consistent set Θ'. Observe that $\Phi'^{C} \subseteq \Theta'$ so that in $M^{CJ'}$, $\Phi' R_C \Theta'$. However $(\Phi', \Theta') \notin \big(\bigcup_{i=1}^{n} R_i\big)^{\mathrm{TC}}$ as for each m, $E^m p \in \Phi'$, but $\neg p \in \Theta'$. Therefore, $M^{CJ'}$ is not a model of $\mathsf{S4}_n^{CJ}$. $\square$

Essentially, $M^{CJ'}$ fails to be an appropriate model because $Cp \not\to Ip$, where I is iterated knowledge.

Filtrations: the general modal case

Filtration is an established technique for producing a finite model from an infinite one so that validity of subformulas is maintained. As in M^{CJ} there are already only a finite number of R_i, a finite model must be one in which W is finite. Each world in the finite model will be an equivalency class of worlds in the original model. We look first at a general modal case, where our modality is '$\Box$.' In the following section we apply these techniques to $M^{CJ'}$ to produce finite counter-models to those formulas not provable in $\mathsf{S4}_n^{CJ}$, concluding the proof of completeness.

Definition 16. For a given set of formulas Φ, say two worlds in a model M are equivalent if they agree on all formulas in Φ:

$$s \equiv_\Phi t \text{ iff } (\forall \psi \in \Phi)(M, s \Vdash \psi \Leftrightarrow M, t \Vdash \psi)$$

and define an equivalence class of worlds

$$[s]_\Phi := \{t \mid s \equiv_\Phi t\} \ ,$$

or simply $[s]$ if Φ is clear.

Note that $\equiv_\Phi$ is indeed an equivalence relation.

Definition 17. A model $N = \langle S, T_1, \dots, T_n, \Vdash_N \rangle$ is a *filtration of M through Φ* if M is a model $\langle W, R_1, \dots, R_n, \Vdash \rangle$ and the following hold:

- Φ is a finite set of formulas closed under subformulas;
- $S = \{[w] \mid w \in W\}$, which is finite as Φ is finite;
- $w \Vdash p \Leftrightarrow [w] \Vdash_N p$ for $p \in \mathit{Var} \cap \Phi$ and $\Vdash_N$ is extended to all formulas;
- Each relation T_i satisfies the following two properties for all modals $\Box$:

$\mathit{min}(T_i/R_i) : (\forall [s], [t] \in S)(\text{if } s' R_i t',\ s' \in [s],\ \text{and }\ t' \in [t],\ \text{then }\ [s] T_i [t])$

$\mathit{max}(T_i/R_i) : (\forall [s], [t] \in S)(\text{if } [s] T_i [t],\ \text{then}$

$$(\forall \Box\psi \in \Phi)[M, s \Vdash \Box\psi \Rightarrow M, t \Vdash \psi]) \ .$$

The condition $min(T_i/R_i)$ ensures that T_i simulates R_i while $max(T_i/R_i)$ permits adding pairs to T_i independently of R_i if it respects $\Box$. Note that a filtration will always exist as you can define the T_i by reconsidering either condition as a bi-implication. This will give the smallest and largest (not necessarily distinct) filtrations, respectively ([BdRV01]).

Theorem 8. *Let N be a filtration of M through Φ, then*

$$(\forall\psi \in \Phi)(\forall s \in W)(M, s \Vdash \psi \Leftrightarrow N, [s] \Vdash_N \psi) \ .$$

Proof. By induction on the complexity of $\psi \in \Phi$.

- $\psi = p$: by definition of $\Vdash_N$.
- I.H.: As Φ closed under subformulas, $M, s \Vdash \psi \Leftrightarrow N, [s] \Vdash_N \psi$ holds for ψ of lower complexity.
- $\psi = \neg\varphi$: $M, s \Vdash \neg\varphi \Leftrightarrow M, s \nVdash \varphi \Leftrightarrow$ (by I.H.) $N, [s] \nVdash_N \varphi \Leftrightarrow N, [s] \Vdash_N \neg\varphi$.
- $\psi = \varphi \wedge \varphi'$: $M, s \Vdash_M \varphi \wedge \varphi' \Leftrightarrow M, s \Vdash_M \varphi$ and $M, s \Vdash_M \varphi' \Leftrightarrow$ (by I.H.) $N, [s] \Vdash_N \varphi$ and $N, [s] \Vdash_N \varphi' \Leftrightarrow N, [s] \Vdash_N \varphi \wedge \varphi'$.

- $\psi = \Box\varphi$: ($\Rightarrow$) Suppose $M, s \Vdash \Box\varphi$. Let $[t]$ be such that $[s]T[t]$. By $max(T/R)$, $M, t \Vdash \varphi$. By I.H. $N, [t] \Vdash_N \varphi$. As $[t]$ was arbitrary, $N, [s] \Vdash_n \Box\varphi$.

  ($\Leftarrow$) Suppose $N, [s] \Vdash_N \Box\varphi$, so $\forall[t]$ such that $[s]T[t]$, $N, t \Vdash_N \varphi$. Let $u \in W$ be any state such that sRu, then by $min(T/R)$ $[s]T[u]$ so that $N, [u] \Vdash_N \varphi$. By I.H. $M, u \Vdash \varphi$ and since u was an arbitrary world accessible from s, $M, s \Vdash \Box\varphi$. $\Box$

Filtrations: the canonical structure $M^{CJ'}$ case

We now consider filtrations in the context of $\mathsf{S4}_n^{CJ}$.

Definition 18. A formula φ has a *suitable set* of subformulas Φ if $\Phi = \Phi_1 \cup \Phi_2 \cup \Phi_3 \cup \Phi_4$ where for $i \in \{1, \ldots, n\}$:

$$\Phi_1 = \{\psi, \neg\psi \mid \psi \text{ is a subformula of } \varphi\} \ ;$$

$$\Phi_2 = \{K_i K_i \psi, \neg K_i K_i \psi \mid K_i\psi \in \Phi_1\} \ ;$$

$$\Phi_3 = \{K_i J\psi, \neg K_i J\psi, K_i\psi, \neg K_i\psi \mid J\psi \in \Phi_1\} \ ;$$

$$\Phi_4 = \{K_i C\psi, \neg K_i C\psi, K_i\psi, \neg K_i\psi \mid C\psi \in \Phi_1\} \ .$$

Crucially, a suitable set is finite and closed under subformulas.

Corollary 7. *Let Φ be a suitable set for φ and M a model such that $M, s \Vdash \varphi$. If N is a filtration of M through Φ, then $N, [s] \Vdash_N \varphi$.*

Proof. By Theorem 8 and $\varphi \in \Phi$. $\square$

Definition 19. For $M^{CJ'} = \langle W, R_1, \ldots, R_n, R_C, R_J, \Vdash \rangle$, the canonical structure of $\mathsf{S4}_n^{CJ}$, and a suitable set Φ for a consistent formula φ, define a model $N = \langle S, T_1, \ldots, T_n, T_C, T_J, \Vdash_N \rangle$ such that, for $i \in \{1, 2, \ldots, n\}$:

- $S = \{[w] \mid w \in W\}$, which is finite as Φ is finite;
- $w \Vdash p \Leftrightarrow [w] \Vdash_N p$ for $p \in \mathit{Var} \cap \Phi$ and $\Vdash_N$ is extended to all formulas;
- $T_i \subseteq S \times S$ such that $[s]T_i[t]$ iff $(s \Vdash K_i\psi \Rightarrow t \Vdash \psi)$ for those $K_i\psi \in \Phi$;
- $T_C = \big(\bigcup_{i=1}^{n} T_i\big)^{\mathrm{TC}}$;
- $T_J \subseteq S \times S$ such that $[s]T_J[t]$ iff $(s \Vdash J\psi \Rightarrow t \Vdash \psi)$ for those $J\psi \in \Phi$.

We now drop the subscript on $\Vdash_N$ to simplify notation. As worlds in N are equivalency classes, it will be clear as to which model is in question.

Lemma 7. *N is a model of $\mathsf{S4}_n^{CJ}$ (see Definition 14).*

Proof. All accessibility relations are reflexive and transitive and $T_i \subseteq T_C \subseteq T_J$.

- T_i is reflexive: For an arbitrary $s \in [s]$, $(s \Vdash K_i\psi \Rightarrow s \Vdash \psi)$ always holds. If the antecedent is true, then by the definition of $\Vdash$ and the reflexivity of R_i, the consequence follows. If the antecedent fails, the implication is vacuously true. Thus for all $[s] \in S$, $[s]T_i[s]$, so T_i is reflexive.

- T_i is transitive: Suppose $[s]T_i[t]$ and $[t]T_i[u]$ and $s \Vdash K_i\psi$ for $K_i\psi \in \Phi$. As Φ is suitable, also $K_iK_i\psi \in \Phi$. As R_i is transitive, the $\mathsf{4}$ axiom is sound so we have $(s \Vdash K_i\psi \Rightarrow s \Vdash K_iK_i\psi)$, so as $[s]T_i[t]$ and $K_iK_i\psi \in \Phi$, $(s \Vdash K_iK_i\psi \Rightarrow t \Vdash K_i\psi)$. Since $K_i\psi \in \Phi$ and $[t]T_i[u]$, $(t \Vdash K_i\psi \Rightarrow u \Vdash \psi)$ so $u \Vdash \psi$. Thus for $K_i\psi \in \Phi$, $(s \Vdash K_i\psi \Rightarrow u \Vdash \psi)$ holds so $[s]T_i[u]$, hence T_i is transitive.

- T_C is reflexive as for every $[s] \in S$, $[s]T_i[s]$ and $T_i \subseteq T_C$. T_C is transitive by definition.

- T_J is reflexive and transitive by the same reasoning as for T_i. It must also be shown that $T_C \subseteq T_J$. Suppose $[s]T_i[t]$, then we want to show $[s]T_J[t]$, i.e. for $J\psi \in \Phi$, $(s \Vdash J\psi \Rightarrow t \Vdash \psi)$ holds. If $J\psi \in \Phi$, then as Φ is suitable, $K_iJ\psi \in \Phi$. Suppose $s \Vdash J\varphi$, then as $M^{CJ'}$ is sound and $\mathsf{S4}_n^{CJ} \vdash J\varphi \to K_iJ\varphi$, $s \Vdash K_iJ\varphi$. Then since $[s]T_i[t]$, $(s \Vdash K_iJ\psi \Rightarrow t \Vdash J\psi)$ holds, so $t \Vdash J\varphi$ holds, and since R_J is reflexive, $t \Vdash \psi$. Thus for $J\psi \in \Phi$ and $[s]T_i[t]$, $(s \Vdash J\psi \Rightarrow t \Vdash \psi)$ holds, so $[s]T_J[t]T$. Since $T_i \subseteq T_J$, $T_C \subseteq T_J$. $\square$

Lemma 8 (Definability Lemma). *Let* $S = \{[s] \mid s \in W\}$ *for some suitable set* Φ*. Then for each subset* $D \subseteq S$ *there is some characteristic formula* χ_D *such that for all* $[s] \in S$*,* $s \Vdash \chi_D$ iff $[s] \in D$ *. Note that all* D *are finite as* S *is.*

Proof. Let the set $\bigwedge\{s\}$ be the conjunction of all $\psi \in \Phi$ that are true at s. By definition of $[s]$, $t \Vdash \bigwedge\{s\}$ iff $[s] = [t]$. Let $\chi_D = \bigvee_{[t]\in D} (\bigwedge\{s\})$.

$$s \Vdash \chi_D \Leftrightarrow \; s \Vdash \bigvee_{[t]\in D} \left(\bigwedge\{s\}\right) \Leftrightarrow s \Vdash \bigwedge\{t\} \text{ for some } t \in [t] \in D$$

$$\Leftrightarrow [s] = [t] \text{ for some } t \in [t] \in D \Leftrightarrow [s] \in D\ .$$

□

Theorem 9. N *of Definition 19 is a filtration of* $M^{CJ'}$ *through* Φ *(cf. [MvdH95]).*

A relation T is a *filtration of* R if it satisfies $min(T/R)$ and $max(T/R)$.

Proof. It needs only to be confirmed that the accessibility relations T_i, T_C, and T_J meet the conditions $min(T/R)$ and $max(T/R)$.

- T_i : T_i satisfies $max(T_i/R_i)$ by definition so it remains to check $min(T_i/R_i)$.

  Suppose $[s], [t] \in S$ with $s' \in [s]$ and $t' \in [t]$ such that $s'R_it'$. For

$K_i\psi \in \Phi$ we have

$$s \vdash K_i\psi \Leftrightarrow s' \vdash K_i\psi \Rightarrow t' \vdash \psi \Rightarrow t \vdash \psi \ .$$

Thus $[s]T_i[t]$ by definition, satisfying $min(T_i/R_i)$.

- T_J : T_J is a filtration of R_J by the same reasoning as in the T_i case, thus $min(T_J/R_J)$ and $max(T_J/R_J)$ are satisfied.

- T_C : To see that T_C satisfies $min(T_C/R_C)$, suppose that sR_Ct. Let $D = \{[w] \in S \mid [s]T_C[w]\}$, the set of worlds reachable from $[s]$ by T_C. It is sufficient to show

$$s \Vdash C\chi_D \ , \tag{2.2}$$

as then sR_Ct gives $t \Vdash \chi_D$ and so by definition of χ_D, $[t] \in D$ and so $[s]T_C[t]$. Now we show (2.2).

As IA is valid in the canonical structure,

$$s \Vdash C(\chi_D \to E\chi_D) \to (\chi_D \to C\chi_D) \ . \tag{2.3}$$

To see that $s \Vdash C(\chi_D \to E\chi_D)$ holds, consider the following. Suppose for some w, sR_Cw and $w \Vdash \chi_D$. We want to show $w \Vdash E\chi_D$, i.e., for all i, $w \Vdash K_i\chi_D$, i.e. for all u, wR_iu, $u \Vdash \chi_D$. Since $w \Vdash \chi_D$, $[w] \in D$ so $[s]T_C[w]$. This means there is a path of length l from $[s]$ to $[w]$ along

> the union of T_is. As each T_i is a filtration of R_i we also have for all those worlds u accessible from w, $[w]T_i[u]$. Thus there is a path of length $l+1$ along the T_is from $[s]$ to $[u]$ and so $[s]T_C[u]$. This means that $u \Vdash \chi_D$, so $w \Vdash E\chi_D$. Since the antecedent of (2.3) holds, we have $s \Vdash \chi_D \rightarrow C\chi_D$ so in order to conclude (2.2), we must show $s \Vdash \chi_D$. which we have by the reflexivity of T_C. Thus T_C satisfies $min(T_C/R_C)$.
>
> T_C must also satisfy $max(T_C/R_C)$. Suppose that $[s]T_C[t]$ and for some $s \in [s]$, $s \Vdash C\psi$ for $C\psi \in \Phi$. We must show that $t \Vdash \psi$. Note that as Φ is suitable, for each i, $K_iC\psi$, i.e. $EC\psi \in \Phi$ as well. Recall from Proposition 7 that $\mathsf{S4}_n^{CJ} \vdash C\psi \rightarrow EC\psi$ so by soundness, $s \Vdash C\psi \Rightarrow s \Vdash EC\psi$. As $[s]T_C[t]$ and T_C is built from filtrations of the R_is, there is a path of length l along the R_is from s to t. As $s \Vdash EC\psi$ and $EC\psi \in \Phi$, $C\psi$ also holds at the next world on this path towards t, for whichever R_i used. By induction on the length of the path we get $t \Vdash C\psi$. Since T_C is reflexive we have $t \Vdash \psi$. Thus T_C satisfies $max(T_C/R_C)$. □

We can now finish the proof of Theorem 7 that $\mathsf{S4}_n^{CJ}$ is sound and complete with respect to $\mathsf{S4}_n^{CJ}$-models. Soundness was shown already.

Proof of Completeness. Suppose $\mathsf{S4}_n^{CJ} \nvdash \varphi$. Then $\{\neg\varphi\}$ is contained in some maximal consistent set Θ and for the canonical structure M' we have $M', \Theta \Vdash \neg\varphi$. Defining a suitable set Φ of subformulas of $\neg\varphi$, we can construct an $\mathsf{S4}_n^{CJ}$-model N (Lemma 7), which, as it happens to be a filtration of M' through Φ (Theorem 9), agrees with M' on formulas of Φ (Theorem 8) and so $N, [\Theta] \nVdash \varphi$. □

Corollary 8. *$\mathsf{S4}_n^{CJ}$ exhibits the Finite Model Property and is decidable.*

Soundness yields the following two propositions.

Proposition 8. *$\mathsf{S4}_n^{CJ} \nvdash C\varphi \to J\varphi$, as was promised before Definition 10.*

Proof. Consider a model of $\mathsf{S4}_2^{CJ}$ with $W = \{a, b\}$ such that $R_1 = R_2 = R_C = \{(a,a),(b,b)\}$ and $R_J = \{(a,a),(b,b),(a,b)\}$. Let only $a \Vdash p$ and all other propositional variable fail at both worlds. While $a \Vdash Cp$, $a \nVdash Jp$ so $a \nVdash Cp \to Jp$ so $a \Vdash \neg(C\varphi \to J\varphi)$, so by soundness $\mathsf{S4}_n^{CJ} \nvdash (C\varphi \to J\varphi)$. □

Proposition 9. *$\mathsf{S4}_n^{CJ}$ is a conservative extension of both $\mathsf{S4}_n^{J}$ and $\mathsf{S4}_n^{C}$.*

Proof. Conservativity of $\mathsf{S4}_n^{CJ}$ over $\mathsf{S4}_n^{J}$:

We need to show that for each $\mathsf{S4}_n^{J}$-formula F, if $\mathsf{S4}_n^{CJ} \vdash F$, then $\mathsf{S4}_n^{J} \vdash F$. By contraposition, suppose $\mathsf{S4}_n^{J} \nvdash F$. Then, by the completeness theorem for $\mathsf{S4}_n^{J}$, there is a $\mathsf{S4}_n^{J}$-model M such that F does not hold in M. Now we transform M into an $\mathsf{S4}_n^{CJ}$-model M' by adding the reachability relation R_C.

This can always be done and leaves the other components of M unaltered. Since the modal C does not occur in F, the truth values of F in M and in M' remains unchanged at each world, hence F does not hold in M'. By soundness of $\mathsf{S4}_n^{CJ}$, $\mathsf{S4}_n^{CJ} \not\vdash F$.

Conservativity of $\mathsf{S4}_n^{CJ}$ over $\mathsf{S4}_n^{C}$:

Let G be an $\mathsf{S4}_n^{C}$-formula not derivable in $\mathsf{S4}_n^{C}$. We have to show that G is not derivable in $\mathsf{S4}_n^{CJ}$ either. By completeness of $\mathsf{S4}_n^{C}$ there is an $\mathsf{S4}_n^{C}$-countermodel N for G. Make N into an $\mathsf{S4}_n^{CJ}$-model N' by the addition of R_J as the total relation (alternatively, we could put $R_J = R_C$). As G contains no J, at each world, these models agree on the valuation of G, thus G does not hold in N' either. By soundness, $\mathsf{S4}_n^{CJ} \not\vdash G$. $\square$

Chapter 3

G.C.K. in Justification Logic

The characteristic feature of *G.C.K.* is that it implies, but is not equivalent to, iterated knowledge I. Logics with this type of common knowledge have already been seen ([Lew69, McSHI78, Art06]), but this new terminology introduced in [Art12b] clarifies this distinction. While generic common knowledge can appropriately be applied in many contexts where C has traditionally been used, there is also a technical advantage in that the cut rule can be eliminated.[1]

In the generative justification logic LP, the logic of proofs, knowledge and reasoning are made explicit with proof terms representing evidence for facts and new logic atoms $t:F$ are introduced with the reading 't is (sufficient) evidence for knowing F' or simply 't is a proof of F.'

---

[1] See details in [AJ05] as to why the finitistic cut-elimination in traditional common knowledge systems may be seen as unsatisfactory.

We define a justification logic with multiple knowers and generic common knowledge. As the standard example, we assume that all agents as well as their *G.C.K.* system are confined to LP. We call the resulting system $\mathsf{LP}_n(\mathsf{LP})$ which symbolically indicates n LP-type agents with an LP-type common knowledge evidence system.

Multi-agent justification logic systems were first considered in [Yav06], but without any common knowledge component. Systems with the explicit equivalent of the traditional common knowledge were considered in [BKS11, Buc12]; capturing common knowledge explicitly proved to be a serious technical challenge and the desirable realization theorem has not yet been obtained.

We take the next step from Theorem 4 (Realization Theorem) by offering a realization of the entire *G.C.K.* system $\mathsf{S4}_n^J$ in the corresponding explicit *G.C.K.* system $\mathsf{LP}_n(\mathsf{LP})$. All epistemic operators, not only J, become explicit in such a realization of $\mathsf{S4}_n^J$ ([Ant13]).

3.1 An explicit epistemic system with *G.C.K.*

Here we introduce an explicit generic common knowledge operator into justification logics in the context of a multi-agent logic of explicit justifications to form a logic $\mathsf{LP}_n(\mathsf{LP})$. The "($\mathsf{LP}$)" corresponds to *G.C.K.*

Definition 20. $\mathcal{L}_{\mathsf{LP}_n(\mathsf{LP})}$, the *language of* $\mathsf{LP}_n(\mathsf{LP})$, is an extension of the propositional language:

$$\mathcal{L}_{\mathsf{LP}_n(\mathsf{LP})} := \{\, \mathit{Var}, \mathit{pfVar}, \mathit{pfConst}, \vee, \wedge, \rightarrow, \neg, +, \cdot, !, : \} \ .$$

Var is propositional variables $(p, q, \dots)$. *Justification terms* (Tm) are built from *pfVar* and *pfConst*, proof variables $(x, y, z, \dots)$ and proof constants $(c, d, \dots)$, by the grammar

$$t := x \mid c \mid t + t \mid t{\cdot}t \mid {!t} \ .$$

Formulas (Fm) are defined by the grammar, for $i \in \{0, 1, 2, \dots, n\}$,

$$\varphi := p \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \varphi \rightarrow \varphi \mid \neg\varphi \mid t :_i \varphi \ ,$$

where $p \in \mathit{Var}$ and $t \in \mathit{Tm}$.

The formulas $t :_i \varphi$ have the intended reading of ‘t is a justification of φ for agent i.’ Index $i = 0$ is reserved for explicit generic common knowledge, for which we will also use the alternative notation $[t]\varphi$ for better readability.

Definition 21. The *axioms and rules of* $\mathsf{LP}_n(\mathsf{LP})$:

CLASSICAL PROPOSITIONAL LOGIC:

A. axioms of classical propositional logic

R. modus ponens

LP AXIOMS FOR ALL $n+1$ AGENTS, $i \in \{0, 1, 2, \ldots, n\}$:

L1. $t{:}_i(\varphi \to \psi) \to (s{:}_i\varphi \to (t{\cdot}s){:}_i\psi)$

L2. $t{:}_i\varphi \to (t+s){:}_i\varphi$ and $t{:}_i\varphi \to (s+t){:}_i\varphi$

L3. $t{:}_i\varphi \to \varphi$

L4. $t{:}_i\varphi \to !t{:}_i(t{:}_i\varphi)$

CONNECTION PRINCIPLE:

C. $[t]\varphi \to t{:}_i\varphi$.

Term operators mirror properties of justifications: "$\cdot$" is *application* for deduction; "+," *sum*, maintains that justifications are not spoiled by adding (possibly irrelevant) evidence; and "!" is *inspection* and stipulates that justifications themselves are justified. This last operator appears only in justification logics with L4, whose corresponding modal logic contains the modal axiom $\mathsf{4}$ ($\Box\varphi \to \Box\Box\varphi$), as shown in [Art08]. A multitude of justification logics of a single agent corresponding to standard modal logics have been developed ([Art08]). Yavorskaya has investigated versions of LP with two agents in which agents can check each other's proofs ([Yav06, Yav08]).

Definition 22. A *constant specification* $\mathcal{CS}_i$ for each agent, $i \in \{0, 1, \ldots, n\}$

is a set of sentences of type $c :_i A$ where c is a constant and A an axiom of $\mathsf{LP}_n(\mathsf{LP})$. Let

$$\mathcal{CS} = \{\mathcal{CS}_1, \ldots, \mathcal{CS}_n\}$$

and $\mathcal{CS}_0 \subseteq \mathcal{CS}_i$ for all $i \in \{1, 2, \ldots, n\}$. By $\mathsf{LP}_{n,\mathcal{CS}}(\mathsf{LP}_{\mathcal{CS}_0})$ we mean the system with the postulates A, R, L1–L4, C above, plus $\mathcal{CS}_0$ and $\mathcal{CS}_i$ as additional axioms. As formulas in a constant specification are taken as axioms, they themselves may be used to form other formulas in a $\mathcal{CS}$ so that it's possible to have $c :_1 (d :_2 A) \in \mathcal{CS}_1$ if $d :_2 A \in \mathcal{CS}_2$.

The constant specification represents assumptions about proofs of basic postulates that are not further analyzed. If $\mathcal{CS}_i = \emptyset$, agent i is totally skeptical; no formulas are justified. If this is so for all agents, the logic would be denoted $\mathsf{LP}_{n,\{\emptyset\}}(\mathsf{LP}_\emptyset)$. Constant specifications of different types have been studied: schematic, injective, full, etc. and have been defined with various closure properties. See [Art08] for a fuller discussion of constant specifications. The *total constant specification* for any agent, $\mathcal{TCS}_i$, is the union of all possible $\mathcal{CS}_i$. In a $\mathcal{TCS}$, any axiom may be justified by any proof constant, or finite string of proof constants, e.g. $c_0 : A$ and $c_m : (\cdots c_2 : (c_1 : (c_0 : A)))$, with each ':' indexed by any i. Henceforth we will assume each agent's constant specification is total and will abbreviate this to $\mathsf{LP}_n(\mathsf{LP})$.

Definition 23. A *modular model of* $\mathsf{LP}_n(\mathsf{LP})$ is $\mathcal{M} = \langle W, R_0, R_1, \ldots, R_n, *, \Vdash\rangle$ where for $i \in \{0, 1, 2, \ldots, n\}$,

1. - W is a nonempty set;
   - $R_i \subseteq W \times W$ are reflexive and R_0 is the designated accessibility relation for *G.C.K.*;
   - $* : W \times \mathit{Var} \to \{0, 1\}$ and $* : W \times \{0, 1, 2, \ldots, n\} \times \mathit{Tm} \to 2^{\mathit{Fm}}$

   i.e., for each agent i at node u, $*(u, i, t)$ is a set of formulas t justifies. We write $t_u^{*,i}$ for $*(u, i, t)$. We assume that *G.C.K.* evidence is everybody's evidence:

   $$t_u^{*,0} \subseteq t_u^{*,i}, \quad \text{for} \quad i \in \{0, 1, 2, \ldots, n\} \ .$$

2. For each agent i and node u, $*$ is closed under the following conditions:

   $$\begin{aligned} \textit{Application:} \quad & s_u^{*,i} \cdot t_u^{*,i} \subseteq (s \cdot t)_u^{*,i} \\ \textit{Sum:} \quad & s_u^{*,i} \cup t_u^{*,i} \subseteq (s + t)_u^{*,i} \\ \textit{Inspection:} \quad & \{t :_i \varphi \mid \varphi \in (t_u^{*,i})\} \subseteq (!t)_u^{*,i} \end{aligned}$$

   where $s^* \cdot t^* = \{\psi \mid \varphi \to \psi \in s^* \text{ and } \varphi \in t^* \text{ for some } \varphi\}$, the set of formulas resulting from applying modus ponens to implications in s^* whose antecedents are in t^*.

3. For $p \in \mathit{Var}$, we define forcing $\Vdash$ for atomic formulas p at node u as

$u \Vdash p$ if and only if $*(u,p) = 1$. To define the truth value of all formulas, extend forcing $\Vdash$ to compound formulas by Boolean laws, and define

$$u \Vdash t{:}_i\,\varphi \quad \Leftrightarrow \quad \varphi \in t_u^{*,i} \ .$$

4. 'justification yields belief' (JYB), i.e., for $i \in \{0,1,2,\ldots,n\}$, $u \Vdash t{:}_i\,\varphi$ yields $v \Vdash \varphi$ for all v such that uR_iv.

Modular models, first introduced for the most basic justification logic in [Art12a], are useful for their clear semantical interpretation of justifications as sets of formulas. For modular models of some other justification logics, refer to [KS12]. For a detailed discussion of the relationship between modular models and Mkrtychev–Fitting models for justification logics, see [Art12a].

A model *respects* $\mathcal{CS}_0,\ldots,\mathcal{CS}_n$ if each $c{:}_i\,A$ in these constant specifications holds (at each world u) in the model.

Theorem 10. $\mathsf{LP}_{n,\mathcal{CS}}(\mathsf{LP}_{\mathcal{CS}_0})$ *is sound and complete. For* $i \in \{0,1,2,\ldots,n\}$*:*

$$\mathsf{LP}_{n,\mathcal{CS}}(\mathsf{LP}_{\mathcal{CS}_0}) \vdash F \text{ iff } F \text{ holds in any modular model respecting } \mathcal{CS}_i \ .$$

Proof of Soundness. By induction on F, for $i \in \{0,1,2,\ldots,n\}$.

- constant specifications: If $c{:}_i\,\varphi \in \mathcal{CS}_i$, then $u \Vdash c{:}_i\,\varphi$ as the model respects $\mathcal{CS}_i$.

- Boolean connectives: hold by definition of the truth of formulas.
- application: Suppose $u \Vdash s :_i (F \to G)$ and $u \Vdash t :_i F$. Then by assumption, $(F \to G) \in s_u^{*,i}$ and $F \in t_u^{*,i}$. Then $G \in s_u^{*,i} \cdot t_u^{*,i} \subseteq (s \cdot t)_u^{*,i}$; thus $u \Vdash (s \cdot t) :_i G$.
- sum: Suppose $u \Vdash t :_i F$. Then $F \in t_u^{*,i}$ and so $F \in s_u^{*,i} \cup t_u^{*,i} \subseteq (s+t)_u^{*,i}$. Thus $u \Vdash (s+t) :_i F$. Likewise, $u \Vdash (t+s) :_i F$.
- modus ponens: Suppose $u \Vdash F \to G$. Then by the definition of the connectives either $u \nVdash F$ or $u \Vdash G$. So if also $u \Vdash F$, then $u \Vdash G$.
- factivity: Suppose $u \Vdash t :_i F$. By the 'justification yields belief' condition, $v \Vdash F$ for all v such that uR_iv. As each R_i is reflexive, uR_iu, so also $u \Vdash F$.
- inspection: Suppose $u \Vdash t :_i F$. Then $F \in t_u^{*,i}$ so $t :_i F \in (!t)_u^{*,i}$. Thus $u \Vdash !t :_i (t :_i F)$.
- connection principle: Suppose $u \Vdash t :_0 F$. Then $F \in t_u^{*,0} \subseteq t_u^{*,i}$ so $u \Vdash t :_i F$. □

Proof of Completeness. By the maximal consistent set construction. For $i \in \{0, 1, 2, \ldots, n\}$, let

- W the set of all maximal consistent sets;
- $\Gamma R_i \Delta$ iff $\Gamma^{i,\#} \subseteq \Delta$ where $\Gamma^{i,\#} = \{F \mid t{:}_i F \in \Gamma\}$;
- For $p \in \mathit{Var}$, $*(\Gamma, p) = 1$ iff $p \in \Gamma$;
- $t^{*,i}_{\Gamma} = \{F \mid t{:}_i F \in \Gamma\}$ (i.e., for $X = p, t{:}_i F$, $\Gamma \Vdash X$ iff $X \in \Gamma$) . □

To confirm that these comprise a modular model, the R_i need to be reflexive, the *G.C.K.* and closure conditions must be checked, and the model must satisfy 'justification yields belief.' As each world is maximally consistent $\Gamma^{i,\#} \subseteq \Gamma$, hence $\Gamma R_i \Gamma$ by L3, so each R_i is reflexive. The *G.C.K.* conditions $t^{*,0}_{\Gamma} \subseteq t^{*,i}_{\Gamma}$ for $i \in \{0, 1, 2, \ldots, n\}$ follow from the C axiom $t{:}_0 F \to t{:}_i F$ for $i \in \{1, 2, \ldots, n\}$. Closure conditions for $\cdot$, $+$, and $!$ follow straightforwardly from the axioms L1, L2, and L4. It remains to check the JYB condition, following the Truth Lemma.

Lemma 9 (Truth Lemma)**.** *$\Gamma \Vdash X$ iff $X \in \Gamma$, for each Γ and X.*

Proof. Induction on X. The atomic and Boolean cases are standard. The only interesting case is $X = t{:}_i F$. Note that $\Gamma \Vdash t{:}_i F$ iff $F \in t^{*,i}_{\Gamma}$ by the definition of modular models. Moreover, under the evaluation particular to this model, $F \in t^{*,i}_{\Gamma}$ iff $t{:}_i F \in \Gamma$. Thus $\Gamma \Vdash t{:}_i F$ iff $t{:}_i F \in \Gamma$. □

Now to see the JYB condition, suppose $\Gamma \Vdash t :_i F$ and consider an arbitrary Δ such that $\Gamma R_i \Delta$. By the definition of this model, $t :_i F \in \Gamma$, hence $F \in \Gamma^{i,\#}$, hence $F \in \Delta$. By the Truth Lemma, $\Delta \Vdash F$.

To finish the proof of completeness, let $\mathsf{LP}_{n,\mathcal{CS}}(\mathsf{LP}_{\mathcal{CS}_0}) \nvdash G$, hence $\{\neg G\}$ is consistent and has a maximal consistent extension, Φ. Since $G \notin \Phi$, by the Truth Lemma, $\Phi \nVdash G$.

Corollary 9. *The canonical model of the completeness proof is transitive.*

Proof. Suppose $\Gamma R_i \Delta$ and $\Delta R_i \Theta$. If $t :_i F \in \Gamma$, then $!t :_i (t :_i F) \in \Gamma$ as Γ is maximal consistent. As $!t :_i (t :_i F) \in \Gamma$ and $\Gamma R_i \Delta$, by the definition of the R_i, $t :_i F \in \Delta$. As $t :_i F \in \Delta$ and $\Delta R_i \Theta$, $F \in \Theta$. Thus if $t :_i F \in \Gamma$, then $F \in \Theta$ that is, $\Gamma R_i \Theta$, hence R_i is transitive in the model of the completeness proof. $\square$

Corollary 10. *Modular models for* LP *(i.e.,* $\mathsf{LP}_0(\mathsf{LP})$*) are* $M = \langle W, R, *, \Vdash \rangle$ *where*

1. - *W is nonempty;*
   - *R is reflexive;*
   - $* : W \times Var \rightarrow \{0,1\}$, $* : W \times Tm \rightarrow 2^{Fm}$.

2. *$*$ closure conditions for $\cdot$, $+$, and $!$.*

3. *$u \Vdash p \Leftrightarrow *(u,p) = 1$ and forcing $\Vdash$ extends a truth value to all formulas by Boolean laws and $u \Vdash t{:}F \Leftrightarrow F \in t^*_u$.*

4. *justification yields belief (JYB): $u \Vdash t : F$ yields $v \Vdash F$ for all v such that uRv .*

These modular models for LP differ from those by presented in [KS12] as no transitivity is required of R, which enlarges the class of modular models for LP. Artemov suggests (personal communication) this modular model for $LP_\emptyset$ which satisfies Definition 4 and is not transitive and hence ruled out by the formulation offered in [KS12]:

- $W = \{a, b, c\}$;
- $R = \{(aa), (bb), (cc), (ab), (bc)\}$;
- $*$ is arbitrary on propositional variables, t^*_a, t^*_b, t^*_c are all empty.

Of course, one could produce more elaborate examples as well, e.g., on the same non-transitive frame, fix a propositional variable p and have $t_1 : t_2 : \ldots : t_n : p$ hold for all proof terms $t_1, \ldots, t_n$, for all n, at any node (in particular, make p true at a, b, c).

While it does not appear to be warranted to confine consideration a priori to transitive modular models, the exact role of transitivity of accessibility relations in modular models is still awaiting a careful analysis.

3.2 Realizing $\mathsf{S4}_n^J$ in $\mathsf{LP}_n(\mathsf{LP})$

We show that $\mathsf{LP}_n(\mathsf{LP})$, a logic of explicit knowledge using proof terms, has a precise modal analog in $\mathsf{S4}_n^J$, an epistemic logic with *G.C.K.*. Recall that J is indeed generic as $J\varphi \to C\varphi$ while $C\varphi \not\to J\varphi$, as illustrated by Propositions 5 and 8.

In Kripke models, a distinction between generic and conventional common knowledge is clear. The accessibility relation for C, R_C, is the exact transitive closure of the union of all agents' accessibility relations R_i. On the other hand, R_J, the relation for J, is any transitive reflexive relation which *contains* the union of all other agents' relations, thus

$$R_{GCK} = R_J \supseteq R_C \ .$$

Speaking generally, there may be flexibility in choosing R_J while R_C is uniquely determined by the agents' R_is. This is paralleled in the explicit case where we also have options for generic common knowledge: there may be many evaluations $*$ such that $t_u^{*,0}$ satisfies $t_u^{*,0} \subseteq t_u^{*,i}$ for all i.

We now have $\mathsf{LP}_n(\mathsf{LP})$ and $\mathsf{S4}_n^J$, each is a multi-agent epistemic logic with generic common knowledge, where all justifications are explicit in the former and implicit in the latter. By proving Theorem 12, we will establish that $\mathsf{LP}_n(\mathsf{LP})$ is the exact explicit version of $\mathsf{S4}_n^J$.

Definition 24. The *forgetful projection* is a translation

$$\circ : \mathcal{L}_{\mathsf{LP}_n(\mathsf{LP})} \to \mathcal{L}_{\mathsf{S4}_n^J}$$

defined inductively as follows:

- $p^\circ = p$, for $p \in \mathit{Var}$;
- $(\neg\psi)^\circ = \neg(\psi^\circ)$;
- $\circ$ commutes with binary Boolean connectives: $(\psi \wedge \varphi)^\circ = \psi^\circ \wedge \varphi^\circ$ and $(\psi \vee \varphi)^\circ = \psi^\circ \vee \varphi^\circ$;
- $(t :_i \psi)^\circ = K_i(\psi^\circ)$ for $i \in \{0, 1, \ldots, n\}$.

Proposition 10. $[\mathsf{LP}_n(\mathsf{LP})]^\circ \subseteq \mathsf{S4}_n^J$.

Proof. The $\circ$ translations of all the $\mathsf{LP}_n(\mathsf{LP})$ axioms and rules are easily seen to be theorems of $\mathsf{S4}_n^J$. $\square$

We want to show that these two logics are really correspondences and that

$$\mathsf{S4}_n^J \subseteq [\mathsf{LP}_n(\mathsf{LP})]^\circ$$

also holds. This is much more involved. Theorem 12 shows that a derivation of any $\mathsf{S4}_2^J$ theorem σ can yield an $\mathsf{LP}_2(\mathsf{LP})$ theorem τ such that $\tau^\circ = \sigma$. This process, the converse of the $\circ$-translation, is a *Realization* r.

Definition 25. A realization r is *normal* if all negative occurrences of modalities, whether a K_i or J, are realized by distinct proof variables.

To provide an algorithm r for such a process, we first give the Gentzen system for $\mathsf{S4}_n^J$ and the Lifting Lemma (Proposition 11).

Definition 26. $\mathsf{S4}_n^J\mathsf{G}$, the *Gentzen version of* $\mathsf{S4}_n^J$, is the usual propositional Gentzen system (e.g. `G1c` in [TS96]) with addition of $n+1$ pairs of rules, where $\Box$ is J or some K_i:

$$\frac{\varphi, \Gamma \ \Rightarrow\ \Delta}{\Box\varphi, \Gamma \ \Rightarrow\ \Delta}\,(\Box,\Rightarrow) \qquad \textit{and} \qquad \frac{J\Gamma, \Box\Delta \ \Rightarrow\ \varphi}{J\Gamma, \Box\Delta \ \Rightarrow\ \Box\varphi}\,(\Rightarrow,\Box) \quad .$$

As usual, capital letters are multisets and $\Box\{\varphi_1, \ldots, \varphi_n\} = \{\Box\varphi_1, \ldots, \Box\varphi_n\}$. In the special case that Δ is empty, the second rule can be read as

$$\frac{J\Gamma \ \Rightarrow\ \varphi}{J\Gamma \ \Rightarrow\ \Box\varphi}\,(\Rightarrow,\Box) \ ,$$

where $\Box$ may be J or some K_i. If instead Γ is empty, the second rule becomes

$$\frac{K_i\Delta \ \Rightarrow\ \varphi}{K_i\Delta \ \Rightarrow\ K_i\varphi}\,(\Rightarrow,\Box) \ ,$$

the usual $(\Rightarrow, \Box)$ rule for $\mathsf{S4}_n\mathsf{G}$.

Theorem 11. $\mathsf{S4}_n^J\mathsf{G}$ *is equivalent to* $\mathsf{S4}_n^J$ *and admits cut-elimination.*

Proof. See Theorem 3 in [Art06]. □

Let $\Gamma = \{\gamma_1, \dots, \gamma_m\}$, $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ be finite lists of formulas, and $\vec{y}$, $\vec{z}$ finite lists of proof variables of matching length, respectively. Then $[\vec{y}\,]\Gamma = [y_1]\gamma_1, \cdots, [y_m]\gamma_m$ and $\vec{z}\,{:}_i\,\Sigma = z_1\,{:}_i\,\sigma_1, \cdots, z_n\,{:}_i\,\sigma_n$, $i \in \{0, 1, 2, \dots, n\}$.

Proposition 11 (Lifting Lemma)**.** *In* $\mathsf{LP}_n(\mathsf{LP})$*, for* $i \in \{0, 1, 2, \dots, n\}$ *and each* $\Gamma, \Sigma, \vec{y}, \vec{z}$*,*

$$\frac{[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash \varphi}{[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash f(\vec{y}, \vec{z})\,{:}_i\,\varphi}$$

for the corresponding proof term $f(\vec{y}, \vec{z})$*.*

In the special case where Σ is empty, we have, for $i \in \{0, 1, 2, \dots, n\}$,

$$\frac{[\vec{y}\,]\Gamma \vdash \varphi}{[\vec{y}\,]\Gamma \vdash f(\vec{y})\,{:}_i\,\varphi} \quad ,$$

where as if Γ is empty, we have, for $i \in \{1, 2, \dots, n\}$ but *not* $i = 0$,

$$\frac{\vec{z}{:}_i\Sigma \vdash \varphi}{\vec{z}{:}_i\Sigma \vdash f(\vec{z})\,{:}_i\,\varphi} \quad .$$

Proof. By induction on the derivation of φ.

- φ is an axiom of $\mathsf{LP}_n(\mathsf{LP})$, then as $\mathsf{LP}_n(\mathsf{LP})$ has $\mathcal{TCS}$, for any constant c, $c\,{:}_i\,\varphi$ so let $f(\vec{y}, \vec{z}) = c$. As $\vdash_{\mathsf{LP}_n(\mathsf{LP})} c\,{:}_i\,\varphi$, also $[\vec{y}\,]\Gamma, \vec{z}\,{:}_i\,\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} c\,{:}_i\,\varphi$. (Here, any $\mathcal{CS}$ in which each axiom has a justification would suffice.)

- φ is $[y_j]\gamma_j$ for some $[y_j]\gamma_j \in [\vec{y}\,]\Gamma$, then $[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} [y_j]\gamma_j$,

  hence $$[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} [!y_j]([y_j]\gamma_j) \ ,$$

  and $$[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} !y_j{:}_i([y_j]\gamma_j) \ .$$

  So $$[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} !y_j{:}_i\varphi \ \text{ for any } i,$$

  so we can put $f(\vec{y}, \vec{z}) = !y_j$.

- φ is $z_j{:}_i\sigma_j$ for some $z_j{:}_i\sigma_j \in \vec{z}{:}_i\Sigma$, then $[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} z_j{:}_i\sigma_j$,

  hence $$[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} !z_j{:}_i(z_j{:}_i\sigma_j) \ .$$

  So $$[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} !z_j{:}_i\varphi \ ,$$

  so let $f(\vec{y}, \vec{z}) = !z_j$.

- φ is derived by modus ponens from ψ and $\psi \to \varphi$. By the induction hypothesis, there exist $t{:}_i\psi$ and $u{:}_i(\psi \to \varphi)$ (where $t = f_t(\vec{y}, \vec{z})$ and $u = f_u(\vec{y}, \vec{z})$). Since $u{:}_i(\psi \to \varphi) \to (t{:}_i\psi \to (u\cdot t){:}_i\varphi)$, by modus ponens $(u\cdot t){:}_i\varphi$. So let $f(\vec{y}, \vec{z}) = (u\cdot t)$.

- φ is $c{:}_i A \in \mathcal{CS}_i$. Since $c{:}_i A \to !c{:}_i(c{:}_i A)$ and $\vdash_{\mathsf{LP}_n(\mathsf{LP})} c{:}_i A$, also $\vdash_{\mathsf{LP}_n(\mathsf{LP})} !c{:}_i(c{:}_i A)$ thus

  $$[\vec{y}\,]\Gamma, \vec{z}{:}_i\Sigma \vdash_{\mathsf{LP}_n(\mathsf{LP})} !c{:}_i\varphi \ .$$

  So let $f(\vec{y}, \vec{z}) = !c$. □

Corollary 11 (Internalization)**.** *Given* $\mathsf{LP}_n(\mathsf{LP}) \vdash \varphi$, *for each* $i \in \{0, 1, \ldots, n\}$

there is some proof term p such that $\mathsf{LP}_n(\mathsf{LP}) \vdash p:_i \varphi$.

Proof. This is a special case of the Lifting Lemma. Each $\mathcal{CS}_i$ is required to contain, for all axioms A, $c:_i A$, for some constant c. These p are *ground terms*, built only from proof constants by application and inspection (and not sum). $\square$

Theorem 12 (Realization Theorem)**.** *If* $\mathsf{S4}_n^J \vdash \varphi$*, then* $\mathsf{LP}_n(\mathsf{LP}) \vdash \varphi^r$ *for some normal realization* r*.*

Though there have now been many realization algorithms presented in the literature for various justification logics, some more efficient, some semantic (e.g. in [Fit05]), this proof follows the structure of the original one from [Art01].

Proof. If $\mathsf{S4}_n^J \vdash \varphi$, then by Theorem 11 there is a cut-free derivation $\mathcal{D}$ of the sequent $\Rightarrow \varphi$ in $\mathsf{S4}_n^J\mathsf{G}$. We now construct a normal realization algorithm r that runs on $\mathcal{D}$ and returns an $\mathsf{LP}_n(\mathsf{LP})$ theorem $\varphi^r = \psi$ such that $\psi^\circ = \varphi$.

In φ, positive and negative modalities are defined as usual. The rules of $\mathsf{S4}_n^J\mathsf{G}$ respect these polarities so that $(\Rightarrow, \square)$ introduces positive occurrences and $(\square, \Rightarrow)$ introduces negative occurrences of $\square$, where $\square$ is J or some K_i. Call the occurrences of $\square$ *related* if they occur in related formulas in the premise and conclusion of some rule: the same formula, that formula boxed or unboxed, enlarged or shrunk by $\wedge$ or $\vee$, or contracted. Extend this notion of related modalities by transitivity. Classes of related $\Box$ occurrences in $\mathcal{D}$ naturally form disjoint *families* of related occurrences. An *essential* family is one in which at least one of its members arises from the $(\Rightarrow, \Box)$ rule, these are clearly positive families.

Now the desired r is constructed by the following three steps so that negative and non-essential positive families are realized by distinct proof variables while essential families will be realized by sums of functions of those proof variables.

Step 1. For each negative family and each non-essential positive family, replace all $\Box$ occurrence so that $J\alpha$ becomes $[x]\alpha$ and $K_i\alpha$ becomes $y :_i \alpha$. Choose new and distinct proof variables x and y for each of these families.

Step 2. Choose an essential family f. Count the number n_f of times the $(\Rightarrow, \Box)$ rule introduces a box to this family. Replace each $\Box$ with a sum of proof terms so that for $i \in \{0, 1, 2, \ldots, n\}$, $K_i\alpha$ becomes

$$(w_1 + w_2 + \cdots + w_{n_f}) :_i \alpha \ ,$$

with each w_j a fresh provisional variable. Do this for each essential family, choosing new provisional variables. The resulting tree $\mathcal{D}'$ is now labeled by $\mathsf{LP}_n(\mathsf{LP})$-formulas.

Step 3. Now the provisional variables need to be replaced, starting with the leaves and working toward the root. By induction on the depth of a node in $\mathcal{D}'$ we will show that after the process passes a node, the sequent at that level becomes derivable in $\mathsf{LP}_n(\mathsf{LP})$ where

$$\Gamma \Rightarrow \Delta$$

is read as provability of

$$\Gamma \vdash_{\mathsf{LP}_n(\mathsf{LP})} \bigvee \Delta \; .$$

Note that axioms $p \Rightarrow p$ and $\bot \Rightarrow$ are derivable in $\mathsf{LP}_n(\mathsf{LP})$. For each move down the tree other than by the rule $(\Rightarrow, \Box)$, the concluding sequent is $\mathsf{LP}_n(\mathsf{LP})$-derivable if its premises are; for rules other that this one, do not change the realization of formulas. For a given essential family f, for the j^{th} occurrence of the $(\Rightarrow, \Box)$ rule, for $\Box$ some K_i, the corresponding node in $\mathcal{D}'$ is labeled

$$\frac{[\vec{z}\,]\Gamma,\ \vec{q}{:_i}\Sigma \Rightarrow \alpha}{[\vec{z}\,]\Gamma,\ \vec{q}{:_i}\Sigma \Rightarrow (u_1 + \cdots + u_{n_f}){:_i}\alpha} \; ,$$

where the zs and qs are proof variables and the us are evidence terms, with u_j a provisional variable.

By the induction hypothesis, the premise is derivable in $\mathsf{LP}_n(\mathsf{LP})$. By the Lifting Lemma (Proposition 11), construct a justification term $f(\vec{z}, \vec{q}\,)$ for α

where

$$[\vec{z}\,]\Gamma,\ \vec{q}{:_i}\Sigma \vdash f(\vec{z},\vec{q}\,){:_i}\alpha\ .$$

Now we will replace the provisional variable u_j as follows

$$[\vec{z}\,]\Gamma,\ \vec{q}{:_i}\Sigma \vdash (u_1 + \cdots + u_{j-1} + f(\vec{z},\vec{q}\,) + u_{j+1} + \cdots + u_{n_f}){:_i}\alpha\ .$$

Substitute each u_j with $f(\vec{z},\vec{q}\,)$ everywhere in $\mathcal{D}'$. There is now one fewer provisional variable in the tree as $f(\vec{z},\vec{q}\,)$ has none. The conclusion to this j^{th} instance of the rule $(\Rightarrow, \Box)$ becomes derivable in $\mathsf{LP}_n(\mathsf{LP})$, completing the induction step.

Eventually all provisional variables are replaced by terms of non-provisional variables, establishing that the root sequent of $\mathcal{D}$, φ^r, is derivable in $\mathsf{LP}_n(\mathsf{LP})$. The realization constructed in this manner is normal. □

Corollary 12. $\mathsf{S4}_n^J$ *is the forgetful projection of* $\mathsf{LP}_n(\mathsf{LP})$.

Proof. A straightforward consequence of Proposition 10 and Theorem 12. □

We see that the common knowledge component of $\mathsf{LP}_n(\mathsf{LP})$ indeed corresponds to the generic common knowledge J and hence can be regarded as the *explicit G.C.K.*

3.3 Realization Example

We demonstrate a realization of an $\mathsf{S4}_2^J$ theorem in $\mathsf{LP}_2(\mathsf{LP})$.

Proposition 12. *$\mathsf{S4}_2^J \vdash J\neg\phi \to K_2\neg K_1\phi$.*

Proof. Here is an $\mathsf{S4}_2^J\mathsf{G}$ derivation of the corresponding sequent.

$$
\begin{array}{rcll}
\phi & \Rightarrow & \phi & \\ \hline
K_1\phi & \Rightarrow & \phi & (\Box, \Rightarrow) \\ \hline
\neg\phi, K_1\phi & \Rightarrow & & (\neg, \Rightarrow) \\ \hline
\neg\phi & \Rightarrow & \neg K_1\phi & (\Rightarrow, \neg) \\ \hline
J\neg\phi & \Rightarrow & \neg K_1\phi & (\Box, \Rightarrow) \\ \hline
J\neg\phi & \Rightarrow & K_2\neg K_1\phi & (\Rightarrow, \Box) \\ \hline
 & \Rightarrow & J\neg\phi \to K_2\neg K_1\phi & (\Rightarrow, \to)
\end{array}
$$

□

Now we follow the realization algorithm to end up with an $\mathsf{LP}_2(\mathsf{LP})$ theorem. In the sequent proof, the J in the root is in negative position and all the Js in this derivation are related and form a negative family. The occurrences of the K_1 modality are all related and they, too, form a negative family. The two occurrences of K_2 form an essential positive family with $n_f = 1$ as there is one use of the $(\Rightarrow, \Box)$ rule.

Step 1. Replace all J occurrences with ‘$[x]$’ and K_1 occurrences with ‘$y\,{:_1}$.’

Step 2. Replace all K_2 occurrences with a ‘$w\,{:_2}$,’ with w a provisional variable. Since here $n_f = 1$, a sum is not required. At this stage the derivation tree looks like this, where '$\Rightarrow$' is read as '$\vdash$' in $\mathsf{LP}_2(\mathsf{LP})$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\phi \Rightarrow \phi}{y{:_1}\phi \Rightarrow \phi}\;{\scriptstyle(\Box,\Rightarrow)}}{\neg\phi, y{:_1}\phi \Rightarrow}\;{\scriptstyle(\neg,\Rightarrow)}}{\neg\phi \Rightarrow \neg y{:_1}\phi}\;{\scriptstyle(\Rightarrow,\neg)}}{[x]\neg\phi \Rightarrow \neg y{:_1}\phi}\;{\scriptstyle(\Box,\Rightarrow)}}{[x]\neg\phi \Rightarrow w{:_2}(\neg y{:_1}\phi)}\;{\scriptstyle(\Rightarrow,\Box)}}{\Rightarrow [x]\neg\phi \to w{:_2}(\neg y{:_1}\phi)}\;{\scriptstyle(\Rightarrow,\to)} \quad .$$

Step 3. The one instance of the $(\Rightarrow,\Box)$ rule calls for the Lifting Lemma to replace w with $f(x)$ so that

$$[x]\neg\phi \vdash f(x){:_2}(\neg y{:_1}\phi)$$

in $\mathsf{LP}_2(\mathsf{LP})$. The proof of the Lifting Lemma is constructive and provides a general algorithm of finding such f. To skip some routine computations we will use the trivial special case of Lifting Lemma, Internalization, without specifying g.

Consider the following Hilbert-style derivation in $\mathsf{LP}_2(\mathsf{LP})$, line 7 in particular.

1. $y{:_1}\phi \to \phi$ L3 axiom for agent 1
2. $\neg\phi \to \neg y{:_1}\phi$ from 1. by contraposition
3. $[g](\neg\phi \to \neg y{:_1}\phi)$ for some ground term g
4. $[g](\neg\phi \to \neg y{:_1}\phi) \to ([x]\neg\phi \to [g{\cdot}x]\neg y{:_1}\phi)$ L1 axiom for *G.C.K.*

5. $[x]\neg\phi \to [g\cdot x]\neg y{:}_1\,\phi$ from 3. and 4. by R1
6. $[g\cdot x]\neg y{:}_1\,\phi \to (g\cdot x){:}_2\,\neg y{:}_1\,\phi$ connection principle
7. $[x]\neg\phi \to (g\cdot x){:}_2\,\neg y{:}_1\,\phi$ from 5. and 6.

So, it suffices to put $f(x) = g\cdot x$ where g is a ground proof term from line 3[2] and $[x]\neg\phi$. Note the forgetful projection of the $\mathsf{LP}_2(\mathsf{LP})$ theorem line 7.,

$$[[x]\neg\phi \to (g\cdot x){:}_2\,\neg y{:}_1\,\phi]^\circ = J\neg\phi \to K_2\neg K_1\phi \ ,$$

is the original $\mathsf{S4}_n^J$ theorem which was realized.

---

[2] Here $g = c\cdot d$ where $[d](y{:}_i\,\phi \to \phi)$ and $[c]((y{:}_i\,\phi \to \phi) \to (\neg\phi \to \neg y{:}_i\,\phi))$. The particular constants c and d depend on $\mathcal{CS}_0$.

Chapter 4

Conclusions

The family of justification logics offers a robust and flexible setting in which to investigate explicit reasons for knowing: $t : \varphi$, 'φ is known for reason t,' in contrast to a modal approach in which $\Box\varphi$ or $K\varphi$ represent implicit knowledge of φ, where reasons are not specified. The addition of generic common knowledge opens these systems to numerous epistemic applications ([Ant07, Art10, Art12b]). The Realization Theorem for $\mathsf{S4}_n^J$ presented here allows for all modalities, including generic common knowledge (J), to be made explicit in $\mathsf{LP}_n(\mathsf{LP})$, enabling reasoning to be tracked.

The construction of $\mathsf{LP}_n(\mathsf{LP})$ can serve as a template to build other multi-agent explicit justification logics with $G.C.K.$, even in cases where not all the agents' reasoning may be factive. If other justification logics such as J, JT, and $\mathsf{J4}$ ([Art08]) were augmented with $G.C.K.$ to form the logics $\mathsf{J}_n(\mathsf{LP})$, $\mathsf{JT}_n(\mathsf{LP})$, $\mathsf{J4}_n(\mathsf{LP})$, it is assumed that they, too, would correspond to implicit modal logics, presumably to K_n^J, KT_n^J, $\mathsf{K4}_n^J$, respectively, but this has yet to be shown. There are also justification logics such as $\mathsf{J45}$, $\mathsf{JD45}$, and $\mathsf{JT45}$ to consider, with the negative introspection axiom $(\neg t{:}_i F \rightarrow ?t{:}_i(\neg t{:}_i F))$, whose forgetful projection is the modal $\mathsf{5}$ axiom $\neg K_i F \rightarrow K_i(\neg K_i F)$ ([Art08]). On the modal side, $G.C.K.$ must be strong enough to be considered knowledge (usually $\mathsf{S4}$), but possibly $\mathsf{S5}$. While there are realization theorems for $\mathsf{S5}$, such as Fitting's clean approach in [Fit08a], it may be more delicate to establish correspondences between these multi-agent $G.C.K.$ systems with negative introspection.

In the $\mathsf{LP}_n(\mathsf{LP})$ case presented here, all agent reasoning represents knowledge. While it is useful to track these justifications, in the knowledge domain, each justification is a proof and so yields truth. However, in a belief setting, justifications are not necessarily sufficient to yield truth. In these situations, it may become even more crucial, indeed essential, to track specific evidence in order to analyze its reliability and compare justifications arriving from different sources. Logics of belief with $G.C.K.$ can be constructed. Without factivity (L3), belief, rather than knowledge, is modeled. Investigating multi-agent logics of belief with $G.C.K.$ will likely also yield a rich source of models in which to analyze several traditional epistemic scenarios and may also offer an entry to considering an explicit version of common belief.

Even within the knowledge context, it may also be worthwhile investigating other levels of group knowledge in an explicit setting. In [FHMV95], a hierarchy of group knowledge is presented, from distributed knowledge at the weakest, to "everybody knows," to finite iterative knowledge, and finally to common knowledge. Understanding these from a justification logic standpoint could enrich the field. However, we currently see that generic common knowledge is a useful choice for modeling many epistemic situations and we have here presented what has yet to be shown for conventional common knowledge: that a modal epistemic logic with generic common knowledge can be made fully explicit. This is done through the introduction of the justification logic $\mathsf{LP}_n(\mathsf{LP})$ with explicit *G.C.K.* and the Realization algorithm.

Chapter 5

Appendix: Muddy Children realized

To consider another instance of a realization, we return to the Muddy Children example of §2.2.1 to obtain a theorem of $\mathsf{LP}_2(\mathsf{LP})$ under an explicit context analougus to the modal $\mathsf{S4}_2^J$ context

$$\mathcal{S} = J(A \vee B) \wedge J(K_b A \vee K_b \neg A) \wedge J(\neg K_b B) \wedge J(\neg K_b \neg B) \ .$$

This realization is done "by hand" and not via the algorithm presented in Theorem 12. The $\mathsf{LP}_2(\mathsf{LP})$ context will be

$$\mathcal{T} = [d](A \vee B) \wedge [d'](p{:}A \vee q{:}\neg A) \wedge [d''](\neg x{:}B) \wedge [d'''](\neg y{:}\neg B) \ .$$

Here, as each conjunct is $G.C.K.$, is to be considered an axiom of this system for any proof terms p, q, x, and y. Here ":" corresponds to b's justifications, ";" to a's and $[\cdot]$ to J's. While we are assuming that all constant specifications are total, distinguishing the proof constants can clarify the process of this derivation. We also note that some steps can be eliminated by sensitively adjusting the conjuncts of $\mathcal{T}$. This is precisely because justification logics are not closed under substitution of equivalent formulas; equivalent sentences may, depending on the constant specification, have different proof terms. We consider propositional tautologies, not only propositional axioms, to be axioms of $\mathsf{LP}_2(\mathsf{LP})$.

Claim. *Given $\mathcal{T}$, child a can prove she is muddy: for some proof term t, $\mathcal{T},\ \mathsf{LP}_2(\mathsf{LP}) \vdash t;A$.*

Proof. $\mathcal{T}, \mathsf{LP}_2(\mathsf{LP}) \vdash t;A$ where $t = c_3 \cdot ((c_2 \cdot d') \cdot (g \cdot d''))$ is built from proof constants are as follows: $c_3;(p : A \to A)$, $c_2;((p : A \vee x : \neg A) \to (\neg x : \neg A \to p : A))$, $g;(\neg(c_1 \cdot d) \cdot x : B \to \neg x : \neg A)$, and $d'';(\neg(c_1 \cdot d) \cdot x : B)$. The term g is not a proof constant itself but a ground term provided for by Internalization (Corollary 11).

The ground term g of line 10. is built by a mimicking the same reasoning in the earlier lines. She justifies any axiom with some constant ($\mathcal{TCS}$). If she wants to duplicate b's or J's modus ponens step she first needs to justify the appropriate L1 statement.

This derivation parallels the modal proof in §2.2.1.

1. $(A \vee B) \to (\neg A \to B)$ propositional tautology

2. $[c_1]((A \vee B) \to (\neg A \to B))$ $\mathcal{TCS}$

3. $[d](A \vee B)$ $\mathcal{T}$

4. $[c_1]((A \vee B) \to (\neg A \to B)) \to ([d](A \vee B) \to [c_1 \cdot d](\neg A \to B))$ L1

5. $[c_1 \cdot d](\neg A \to B)$ from 4., 2., and 3. by R twice

6. $c_1 \cdot d : (\neg A \to B)$ from 5. by C

7. $c_1 \cdot d : (\neg A \to B) \to (x : \neg A \to (c_1 \cdot d) \cdot x : B)$ L1

8. $x : \neg A \to (c_1 \cdot d) \cdot x : B$ from 6. and 7. by m.p.

9. $\neg (c_1 \cdot d) \cdot x : B \to \neg (x : \neg A)$ contrapositive of 8.

10. $g ; (\neg (c_1 \cdot d) \cdot x : B \to \neg x : \neg A)$ for g a ground term, see above

11. $[d''](\neg (c_1 \cdot d) \cdot x : B)$ $\mathcal{T}$

12. $d'' ; (\neg (c_1 \cdot d) \cdot x : B)$ from 11. by C

13. $g ; (\neg (c_1 \cdot d) \cdot x : B \to \neg x : \neg A) \to (d'' ; (\neg (c_1 \cdot d) \cdot x : B) \to g \cdot d'' ; (\neg x : \neg A))$ L1

14. $g \cdot d'' ; (\neg x : \neg A)$ from 13., 10., and 12. by R twice

15. $(p : A \vee x : \neg A) \to (\neg x : \neg A \to p : A)$ propositional tautology

16. $[c_2]((p : A \vee x : \neg A) \to (\neg x : \neg A \to p : A))$ $\mathcal{TCS}$

17. $[d'](p : A \vee x : \neg A)$ $\mathcal{T}$

18. $[c_2 \cdot d'](\neg x : \neg A \to p : A)$ from 16., 17., L1, and R

19. $c_2 \cdot d' ; (\neg x : \neg A \to p : A)$ from 18. by C

20. $(c_2 \cdot d') \cdot (g \cdot d'') ; (p : A)$ from 19., 14., L1, and R

21. $p : A \to A$ L3

22. $c_3 ; (p : A \to A)$ $\mathcal{TCS}$

23. $c_3 \cdot ((c_2 \cdot d') \cdot (g \cdot d'')) ; A$ from 22., 20., L1, and R. □

Bibliography

[AJ05] Alberucci, L. & Jaeger, G. (2005). About cut elimination for logics of common knowledge. *Annals of Pure and Applied Logic*, 133(1–3), 73–99. (cit. on pp. 30, 60)

[Ant13] Antonakos, E. (2013). Explicit Generic Common Knowledge. In: Artemov, S., Nerode, A. (Eds.) *Logical Foundations of Computer Science, International Symposium, LFCS 2013, San Diego, CA, USA, January 2013* (pp. 16–28). LNCS, Vol. 7734, Springer. (cit. on pp. 61)

[Ant07] Antonakos, E. (2007). Justified and common knowledge: Limited conservativity. In: Artemov, S., Nerode, A. (Eds.) *Logical Foundations of Computer Science, International Symposium, LFCS 2007, New York, NY, USA, June 4–7, 2007, Proceedings* (pp. 1–11). LNCS, Vol. 4514, Springer. (cit. on pp. 12, 83)

[Art12b] Artemov, S. (2012). Multiplicity of Common Knowledge [PDF document]. Prague Workshop on Non-classical epistemic logics, June 14-15, 2012 Abstracts site. Retrieved from: `http://logika.flu.cas.cz/redaction.php?action=showRedaction&id_categoryNode=1996` . (cit. on pp. 12, 30, 60, 83)

[Art12a] Artemov, S. (2012). The ontology of justifications in the logical setting. *Studia Logica*, 100(1–2), 17–30. (cit. on pp. 66)

[Art10] Artemov, S. (2010). Robust Knowledge and Rationality. Technical Report TR-2010010, CUNY Ph.D. Program in Computer Science. Re-

trieved from: `http://tr.cs.gc.cuny.edu/tr/techreport.php?id=406` .
(cit. on pp. 12, 83)

[Art08] Artemov, S. (2008). The logic of justification. *The Review of Symbolic Logic*, 1(4), 477–513.
(cit. on pp. 63, 64, 83)

[Art06] Artemov, S. (2006). Justified Common Knowledge. *Theoretical Computer Science*, 357(1–3), 4–22.
(cit. on pp. 21, 24, 26, 30, 60, 74)

[Art04] Artemov, S. (2004). Evidence-Based Common Knowledge. Technical Report TR-2004018, CUNY Ph.D. Program in Computer Science. Retrieved from: `http://tr.cs.gc.cuny.edu/tr/techreport.php?id=140` .
(cit. on pp. 15, 21)

[Art01] Artemov, S. (2001). Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, 7(1), 1–36.
(cit. on pp. 13, 14, 30, 76)

[AF11] Artemov, S. & Fitting, M. Justification Logic. In E.N. Zalta (Ed), *The Stanford Encyclopedia of Philosophy (Fall 2012 Edition)*. Retrieved from: `http://plato.stanford.edu/archives/fall2012/entries/logic-justification` .
(cit. on pp. 2)

[Aum76] Aumann, R.J. (1976). Agreeing to Disagree. *Annals of Statistics*, 4(6), 1236–1239.
(cit. on pp. 12, 15, 16)

[Bal10] Ballarin, R. (2010). Modern Origins of Modal Logic. In E.N. Zalta (Ed), *The Stanford Encyclopedia of Philosophy (Winter 2010 Edition)*. Retrieved from: `http://plato.stanford.edu/archives/win2010/entries/logic-modal-origins/` .
(cit. on pp. 13)

[Bar88] Barwise, J. (1988). Three Views of Common Knowledge. In *Proceedings of the 2nd conference on Theoretical aspects of reasoning about*

knowledge TARK '88 (pp. 365-379). San Francisco, CA, USA: Morgan Kaufman Publishers Inc.
(cit. on pp. 12, 17)

[vBS04] van Benthem, J. & Sarenac, D. (2004). The Geometry of Knowledge. ILLC Report PP-2004-21, University of Amsterdam. Retrieved from: `http://www.illc.uva.nl/Research/Reports/PP-2004-20.text.pdf` .
(cit. on pp. 17)

[BdRV01] Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal Logic*, Cambridge Tracts in Theoretical Computer Science, Vol. 53. Cambridge, UK: Cambridge University Press.
(cit. on pp. 7, 51)

[Bry05] Bryukhov, Y (2005). *Integration of decision procedures into high-order interactive provers.* (Doctoral dissertation). CUNY Graduate School, New York, NY, USA.
(cit. on pp. 31)

[Buc12] Bucheli, S. (2012). *Justification Logics with Common Knowledge.* (Doctoral dissertation). Universität Bern, Bern, Switzerland.
(cit. on pp. 61)

[BKS11] Bucheli, S., Kuznets, R., & Studer, T. (2011). Justifications for common knowledge. *Journal of Applied Non-Classical Logics*, 21(1), 35–60.
(cit. on pp. 61)

[Dem00] Demri, S. (2000). Complexity of Simple Dependent Bimodal Logics. In R. Dyckoff (Ed) *Automated Reasoning with Analytic Tableau and Related Methods* (pp. 190–204). LNCS Vol. 1847, Springer.
(cit. on pp. 31)

[FHMV95] Fagin, R., Halpern, J., Moses, Y., & Vardi, M. (1995). *Reasoning About Knowledge.* Cambridge, MA: MIT Press. (1st MIT paperback ed., 2004).
(cit. on pp. 10, 16, 17, 19, 23, 26, 31, 85)

[Fit08a] Fitting, M. (2008). The Realization Theorem for S5: A Simple, Constructive Proof. In J. van Benthem, A. Gupta & E. Pacuit (Eds),

Games, Norms and Reasons, Chapter 4, 61-76. Springer.
(cit. on pp. 84)

[Fit05] Fitting, M. (2005). The Logic of Proofs, Semantically. *Annals of Pure and Applied Logic*, 132: 1–25.
(cit. on pp. 76)

[Gea92] Geanakoplos, J. (1992). Common knowledge. In *Proceedings of the 4th conference on Theoretical aspects of reasoning about knowledge TARK '92*, (pp. 254-315). Morgan Kaufmann Publishers Inc.
(cit. on pp. 17)

[HM90] Halpern J. & Moses, Y. (1990). Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3): 549-587. (A preliminary version appeared in 1984.)
(cit. on pp. 15, 16)

[Hin62] Hintikka, J. (1962). *Knowledge and belief: an introduction to the logic of the two notions.* Ithaca, NY NY: Cornell University Press.
(cit. on pp. 6)

[Kri59] Kripke, S. (1959). A completeness Theorem in Modal Logic. *The Journal of Symbolic Logic*, 24(1), 1–14.
(cit. on pp. 7)

[KS12] Kuznets R., Studer, T, (2012). Justifications, Ontology, and Conservativity. In: Bolander, T., Braüner, T., Ghilardi, S., & Moss, L. (eds.) *Advances in Modal Logic*, Vol. 9, 437–458.
(cit. on pp. 66, 70)

[Leh84] Lehmann, D. (1984). Knowledge, common knowledge, and related puzzles. *Proceeding of the 3rd ACM Symposium on Principles of Distributed Computing*, 62-67.
(cit. on pp. 16)

[Lew69] Lewis, D (1969). *Convention, A Philosophical Study.* Cambridge, MA: Harvard University Press.
(cit. on pp. 15, 16, 60)

[McSHI78] McCarthy, J., Sato, M., Hayashi, T., Igarishi, S. (1978) On the Model Theory of Knowledge. Technical Report STAN-CS-78-657, Stanford University. Retrieved from: `ftp://db.stanford.edu/pub/cstr/reports/cs/tr/78/657/CS-TR-78-657.pdf`
(cit. on pp. 15, 16, 31, 32, 60)

[MvdH95] Meyer, J.-J., & van der Hoek, W. (1995). *Epistemic Logic of AI and Computer Science.* Cambridge Tracts in Theoretical Computer Science, Vol. 41. Cambridge, UK: Cambridge University Press.
(cit. on pp. 18, 19, 47, 55)

[Pla] Plato (1996). *Theaetetus; Sophist* (H.N. Fowler, Trans). Loeb Classics Library, Vol. 123. Cambridge, MA: Harvard University Press. (Originally published 1921).
(cit. on pp. 1)

[Rub06] Rubtsova, N. (2006). On realization of S5 modality by evidence terms. *Journal of Logic and Computation*, 16, 671-684.
(cit. on pp. 26, 30)

[TS96] Troelstra, A.S., Schwittenberg, H. (1996). *Basic Proof Theory.* Cambridge Tracts in Theoretical Computer Science Vol. 43, Cambridge University Press.
(cit. on pp. 73)

[Yav08] Yavorskaya (Sidon), T. (2008). Interacting explicit evidence systems. *Theory of Computing Systems*, 43(2), 272–293. Published online October 2007.
(cit. on pp. 63)

[Yav06] Yavorskaya (Sidon), T. (2006). Multi-agent explicit knowledge. In D. Grigoriev, J. Harrison, & E.A. Hirsch (Eds.), *Computer Science - Theory and Applications, First International Computer Science Symposium in Russia, CSR 2006, St. Petersburg, Russia, June 8-12, 2006, Proceedings*, (pp. 369–380). LNCS, Vol. 3967. Springer.
(cit. on pp. 61, 63)