Reasoning-based introspection*

Olivier Gossner[†]

Paris School of Economics, France

London School of Economics and Political Science, UK

Elias Tsakas[‡]

Maastricht University, The Netherlands

December 20, 2010

Abstract

We show that if an agent reasons according to standard inference rules, the axioms of truth and introspection extend from the set of non-epistemic propositions to the whole set of propositions. This implies that the usual axiomatization of the partitional possibility correspondence, which describes an agent who processes information rationally, is redundant.

Keywords: Knowledge, introspection, truth axiom, partitional information structures, epistemic game theory.
JEL Classification: D80, D83, D89.

1 Introduction

The information of an agent who processes information rationally is commonly represented by a partition over a state space (Aumann, 1976). The usual axiomatization for

*We are indebted to Geir Asheim, Giacomo Bonanno, Amanda Friedenberg, Willemien Kets, Friederike Mengel, Andrés Perea, Mark Voorneveld and Jörgen Weibull for fruitful discussions and useful comments. We would also like to thank the audiences of the Epistemic Game Theory Workshop (Stony Brook), Game Theory Conference (Stony Book), ESEM (Barcelona), SAET (Ischia), CRETE (Tinos), Bocconi University, LSE, Santa Fe Institute, Paris Game Theory Seminar, IHPST (Paris), Stockholm School of Economics, University of Warwick, HEC Lausanne and Göteborg University. The financial support from the Marie Curie Intra-European Fellowship is gratefully acknowledged (PIEF-GA-2009-237614).

[†]E-mail: ogossner@ens.fr
[‡]E-mail: e.tsakas@maastrichtuniversity.nl

this representation is that of an agent whose knowledge satisfies truth and introspection (Fagin et al., 1995; Samet, 1990; Aumann, 1999). The truth axiom says that any proposition known to the agent is true. According to the introspection axioms, the agent knows both what he knows and what he does not know.

It is generally admitted that the agent's capacity to draw logical inferences is part of the properties that define the agent's rationality (e.g., see Geanakoplos, 1989). However, it is not clear why a rational agent should also be endowed with introspective abilities. The objective of this note is to investigate to what extent truth and introspection can be explained by a deductive process on the part of the agent.

We consider an agent who observes natural facts about the surrounding world. These natural facts are those described by non-epistemic – else called Boolean – propositions and correspond to sentences that do not involve the agent's own knowledge. The agent also observes his own knowledge about natural facts, i.e., he knows what he knows and what he does not know about natural facts.

We show that the truth and introspection axioms are then satisfied for every proposition, whether epistemic or not. This result provides a justification for the truth axiom and introspection for the epistemic propositions that is based on reasoning.

We express our results in terms of axiomatizations of knowledge. Formally, we show that the standard axiomatization of syntactic knowledge $S5$ is unchanged if truth and introspection are assumed on non-epistemic propositions only. In this sense, our result shows that the usual axiomatization $S5$ is redundant, as it can be replaced by a smaller set of axioms. It is well known that in $S5$, every proposition is equivalent to a proposition of epistemic depth at most one (an early reference to this result is Halpern, 1995). Nevertheless, our result does not appear to be a consequence of this fact, and our proof, although relatively elementary, cannot be reduced to this remark.

The axiomatic model and main result are presented in Section 2. In Section 3, we recall the connection between $S5$ and partitional models, discuss the tightness of the weaker axiomatization presented, and present a multi-agent extension.

2 Model and main result

We recall the standard syntactic model of knowledge from (Chellas, 1980; Fagin et al., 1995). Let Φ be the alphabet of the agent's language, called the set of atomic propositions. These atomic propositions express basic facts about nature such as "it is raining in New York", or "the cat is mortal". The set of non-epistemic (else called Boolean) propositions $\mathcal{L}^0(\Phi)$, with generic elements ϕ, ψ, is the closure of Φ with respect to the standard connectives of negation, $\neg$, and conjunction, $\wedge$.

The knowledge modality is denoted K, and $K\phi$ stands for "the agent knows ϕ". The set of all propositions $\mathcal{L}(\Phi)$, with generic elements ϕ, ψ, is the closure of Φ with respect to $\neg$, $\wedge$ and the knowledge modality K. The propositions $\phi \vee \psi$, $\phi \to \psi$ and $\phi \leftrightarrow \psi$ stand for $\neg(\neg\phi \wedge \neg\psi)$, $\neg\phi \vee \psi$, and $(\phi \to \psi) \wedge (\psi \to \phi)$ respectively.

We recall the standard Modal Logic (ML) system $S5$, consisting of the following axioms and inference rules:

A_1. All tautologies of propositional calculus

A_2. $\big(K\phi \wedge K(\phi \to \psi)\big) \to K\psi$ (Axiom of distribution)

A_3. $K\phi \to \phi$ (Truth axiom)

A_4. $K\phi \to KK\phi$ (Positive introspection)

A_5. $\neg K\phi \to K\neg K\phi$ (Negative introspection)

R_1. From ϕ and $(\phi \to \psi)$ infer ψ (Modus Ponens)

R_2. From ϕ infer $K\phi$ (Rule of necessitation)

The first axiom, A_1, refers to logical propositions such as $(\phi \to \psi) \leftrightarrow (\neg\psi \to \neg\phi)$, which are always logically true; A_2 says that if the agent knows that ϕ implies ψ, and knows ϕ, then he necessarily knows ψ; the truth axiom says that the agent cannot wrongly know a proposition; positive introspection states that the agent knows what he knows, whereas negative introspection says that the agent knows what he does not know.

The propositions that can be proven from the axioms, or from other propositions that have already been proven, are called theorems. Formally, the set of theorems is the closure of the axioms with respect to the inference rules. Let T_5 denote the set of theorems in $S5$.

We show that $S5$ is unchanged if Truth and Introspection are assumed for natural propositions only. Formally, let $S5^0$ be the system consisting of the axioms:

$A_0 - A_2$ for all propositions in $\mathcal{L}(\Phi)$,

$A_3 - A_5$ for all propositions in $\mathcal{L}^0(\Phi)$, and together with the inference rules R_1 and R_2. Wet let T_5^0 denote the set of theorems in $S5^0$.

Our main theorem shows that it suffices to assume the truth axiom and introspection only for the non-epistemic propositions $\mathcal{L}^0(\Phi)$ in order to obtain the same knowledge as in $S5$.

Main Theorem. $T_5^0 = T_5$.

2.1 Proof of the Main Theorem

The general strategy of the proof is to show that (i) the set of propositions that satisfy truth and introspection is closed under a number of operations, and (ii) these operations are sufficient to generate the whole set of propositions starting from the non-epistemic propositions only. It is relatively straightforward to show that $A_3 - A_5$ for ϕ and $\neg\phi$ imply $A3 - A5$ for $K\phi$ and $\neg K\phi$ (Lemma 2), and that $A_3 - A_5$ for ϕ and ψ imply $A_3 - A_5$ for $\phi \wedge \psi$ (Lemma 3). However, it is not true in general that the set of propositions that satisfy $A_3 - A_5$ is closed under $\vee$. Instead, we show that if ϕ, ψ satisfy truth and introspection, so do $K\phi \vee \psi$ and $\neg K\phi \vee \psi$ (Lemma 5).

Remark 1. The following property, called consistency, states that the agent cannot believe a proposition and its negation simultaneously.

A_0. $K\phi \to \neg K \neg \phi$ (Consistency axiom)

It is a consequence of $A_1 - A_3$, hence is automatically satisfied in $S5^0$. We make repeated use of it in the proofs. $\triangleleft$

We divide the proof of the Main Theorem into a series of lemmata. In the proofs we use the convenient notation $\phi \Rightarrow \psi$ to express that ψ is a theorem deduced from ϕ. We indicate, as a superscript of "$\Rightarrow$", which axioms, rules of inference, or previous lemmata (denoted by $L.$) are used in this deduction.

Definition 1. *For some $\Psi \subseteq \mathcal{L}(\Phi)$, let $S5(\Psi)$ denote the ML system consisting of*

- *the axioms $A_0 - A_2$ for all proposition in $\mathcal{L}(\Phi)$,*
- *the axioms $A_3 - A_5$ for all propositions in Ψ, and*
- *the inference rules $R_1 - R_2$.*

Lemma 1. *$(K\phi \wedge K\psi) \leftrightarrow K(\phi \wedge \psi)$ is a theorem in $S5(\Psi)$ for any $\Psi \subseteq \mathcal{L}(\Phi)$.*

Proof. $(\rightarrow)$: The proposition $\phi \rightarrow \big(\psi \rightarrow (\phi \wedge \psi)\big)$ is a tautology of propositional calculus, and therefore it follows from R_2 and A_2 that $K\phi \rightarrow K\big(\psi \rightarrow (\phi \wedge \psi)\big) \overset{A_1}{\Longrightarrow} K\phi \rightarrow \big(K\psi \rightarrow K(\phi \wedge \psi)\big)$, which can be rewritten as $(K\phi \wedge K\psi) \rightarrow K(\phi \wedge \psi)$.

$(\leftarrow)$: the proposition $(\phi \wedge \psi) \rightarrow \phi$ is a tautology of propositional calculus, and therefore it follows from R_2 and A_2 that $K(\phi \wedge \psi) \rightarrow K\phi$. Likewise, for $(\phi \wedge \psi) \rightarrow \psi$, which completes the proof. □

Lemma 2. *$A_3 - A_5$ for $K\phi$ and $\neg K\phi$ are theorems in $S5(\{\phi, \neg\phi\})$.*

Proof. (A_3 for $K\phi$) : The agent assumes A_3 for ϕ. Hence, $(K\phi \rightarrow \phi) \overset{R_2}{\Longrightarrow} K(K\phi \rightarrow \phi) \overset{A_2,R_1}{\Longrightarrow} (KK\phi \rightarrow K\phi)$.

(A_3 for $\neg K\phi$) : The agent assumes A_0 for $K\phi$. Thus, $(KK\phi \rightarrow \neg K\neg K\phi) \overset{A_4,R_1}{\Longrightarrow} (K\phi \rightarrow \neg K\neg K\phi) \overset{A_1,R_1}{\Longrightarrow} (K\neg K\phi \rightarrow \neg K\phi)$.

(A_4 for $K\phi$) : The agent assumes A_4 for ϕ. Hence, $(K\phi \rightarrow KK\phi) \overset{R_2}{\Longrightarrow} K(K\phi \rightarrow KK\phi) \overset{A_2,R_1}{\Longrightarrow} (KK\phi \rightarrow KKK\phi)$.

(A_4 for $\neg K\phi$) : The agent assumes A_5 for ϕ. Hence, $(\neg K\phi \rightarrow K\neg K\phi) \overset{R_2}{\Longrightarrow} K(\neg K\phi \rightarrow K\neg K\phi) \overset{A_2,R_1}{\Longrightarrow} (K\neg K\phi \rightarrow KK\neg K\phi)$.

(A_5 for $K\phi$) : The agent assumes A_4 for ϕ. Thus, $(K\phi \rightarrow KK\phi) \overset{A_1,R_1}{\Longrightarrow} (\neg KK\phi \rightarrow \neg K\phi) \overset{A_5,R_1}{\Longrightarrow} (\neg KK\phi \rightarrow K\neg K\phi)$. Furthermore, the agent has already proven A_3 for $K\phi$, and therefore, $(KK\phi \rightarrow K\phi) \overset{A_1,R_1}{\Longrightarrow} (\neg K\phi \rightarrow \neg KK\phi) \overset{R_2}{\Longrightarrow} K(\neg K\phi \rightarrow \neg KK\phi) \overset{A_2,R_1}{\Longrightarrow} (K\neg K\phi \rightarrow K\neg KK\phi)$. Combining the two we obtain $(\neg KK\phi \rightarrow K\neg K\phi) \wedge (K\neg K\phi \rightarrow K\neg KK\phi) \overset{R_1}{\Longrightarrow} (\neg KK\phi \rightarrow K\neg KK\phi)$, which completes the proof.

(A_5 for $\neg K\phi$) : The agent assumes A_5 for ϕ. Thus, $(\neg K\phi \rightarrow K\neg K\phi) \overset{A_1,R_1}{\Longrightarrow} (\neg K\neg K\phi \rightarrow K\phi) \overset{A_4,R_1}{\Longrightarrow} (\neg K\neg K\phi \rightarrow KK\phi) \overset{A_4,R_1}{\Longrightarrow} (\neg K\neg K\phi \rightarrow KKK\phi)$. Moreover, the agent assumes A_0 for $K\phi$, implying $(KK\phi \rightarrow \neg K\neg K\phi) \overset{R_2}{\Longrightarrow} K(KK\phi \rightarrow \neg K\neg K\phi) \overset{A_2,R_1}{\Longrightarrow} (KKK\phi \rightarrow K\neg K\neg K\phi)$. Combining the two yields $(\neg K\neg K\phi \rightarrow KKK\phi) \wedge (KKK\phi \rightarrow K\neg K\neg K\phi) \overset{R_1}{\Longrightarrow} (\neg K\neg K\phi \rightarrow K\neg K\neg K\phi)$, which completes the proof. □

Lemma 3. *$A_3 - A_5$ for $\phi \wedge \psi$ are theorems in $S5(\{\phi, \psi\})$.*

Proof. (A_3) It follows from Lemma 1 that $\big(K(\phi \wedge \psi) \rightarrow (K\phi \wedge K\psi)\big)$. Hence, $\big(K(\phi \wedge \psi) \rightarrow (K\phi \wedge K\psi)\big) \overset{A_3,R_1}{\Longrightarrow} \big(K(\phi \wedge \psi) \rightarrow (\phi \wedge \psi)\big)$.

(A_4) It follows from Lemma 1 that $\big(K(\phi \wedge \psi) \rightarrow (K\phi \wedge K\psi)\big)$. Hence, $\big(K(\phi \wedge \psi) \rightarrow (K\phi \wedge K\psi)\big) \overset{A_4,R_1}{\Longrightarrow} \big(K(\phi \wedge \psi) \rightarrow (KK\phi \wedge KK\psi)\big) \overset{L.1,R_1}{\Longrightarrow} \big(K(\phi \wedge \psi) \rightarrow K(K\phi \wedge K\psi)\big) \overset{L.1,R_1}{\Longrightarrow} \big(K(\phi \wedge \psi) \rightarrow KK(\phi \wedge \psi)\big)$.

(A_5) It follows from A_1 and Lemma 1 that $\big(\neg K(\phi \wedge \psi) \rightarrow (\neg K\phi \vee \neg K\psi)\big)$. Hence, $\big(\neg K(\phi \wedge \psi) \rightarrow (\neg K\phi \vee \neg K\psi)\big) \overset{A_5,R_1}{\Longrightarrow} \big(\neg K(\phi \wedge \psi) \rightarrow (K\neg K\phi \vee K\neg K\psi)\big) \overset{A_2,R_1}{\Longrightarrow} \big(\neg K(\phi \wedge$

$\psi) \rightarrow K(\neg K\phi \vee \neg K\psi)\big) \overset{A_1,A_2,R_1}{\Longrightarrow} \big(\neg K(\phi \wedge \psi) \rightarrow K\neg(K\phi \wedge K\psi)\big) \overset{L.1,R_1}{\Longrightarrow} \big(\neg K(\phi \wedge \psi) \rightarrow K\neg K(\phi \wedge \psi)\big)$. $\square$

Lemma 4. *$K(K\phi \vee \psi) \leftrightarrow (K\phi \vee K\psi)$ is a theorem in $S5(\{\phi, \psi\})$.*

Proof. $(\leftarrow)$: The agent assumes A_4 for ϕ. Hence, $\big((K\phi \vee K\psi) \rightarrow (KK\phi \vee K\psi)\big) \overset{A_2,L.1,R_1}{\Longrightarrow} \big((K\phi \vee K\psi) \rightarrow K(K\phi \vee \psi)\big)$.

$(\rightarrow)$: The agent assumes A_1. Therefore, $\big(K(K\phi \vee \psi) \rightarrow K(\neg K\phi \rightarrow \psi)\big) \overset{A_2,R_1}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow (K\neg K\phi \rightarrow K\psi)\big) \overset{A_1}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow (\neg K\neg K\phi \vee K\psi)\big) \overset{A_1,A_5,R_1}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow (K\phi \vee K\psi)\big)$. $\square$

Remark 2. Aumann (1999) obtained the same conclusion as in Lemma 4, having assumed the Truth axiom for $K\phi \vee \psi$, which is not assumed here. $\triangleleft$

Lemma 5. *$A_3 - A_5$ for $K\phi \vee \psi$ and $\neg K\phi \vee \psi$ are theorems in $S5(\{\phi, \psi\})$.*

Proof. (A_3 for $K\phi \vee \psi$) : It follows from Lemma 4 that the agent has proven $K(K\phi \vee \psi) \rightarrow (K\phi \vee K\psi)$, implying $\big(K(K\phi \vee \psi) \rightarrow (K\phi \vee K\psi)\big) \overset{A_3,R_1}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow (K\phi \vee \psi)\big)$.

(A_3 for $\neg K\phi \vee \psi$) : The agent assumes A_5 for ϕ, implying $\big((\neg K\phi \vee \psi) \rightarrow (K\neg K\phi \vee \psi)\big) \overset{R_2}{\Longrightarrow} K\big((\neg K\phi \vee \psi) \rightarrow (K\neg K\phi \vee \psi)\big) \overset{A_2,R_1}{\Longrightarrow} \big(K(\neg K\phi \vee \psi) \rightarrow K(K\neg K\phi \vee \psi)\big) \overset{L.2,L.4,R_1}{\Longrightarrow} \big(K(\neg K\phi \vee \psi) \rightarrow (K\neg K\phi \vee K\psi)\big) \overset{L.2,R_1}{\Longrightarrow} K(\neg K\phi \vee \psi) \rightarrow (\neg K\phi \vee \psi)$.

(A_4 for $K\phi \vee \psi$) : It follows from Lemma 4 that the agent has proven $K(K\phi \vee \psi) \rightarrow (K\phi \vee K\psi)$, implying $\big(K(K\phi \vee \psi) \rightarrow (K\phi \vee K\psi)\big) \overset{L.2,R_I}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow (KKK\phi \vee KK\psi)\big) \overset{A_2,R_1,L.1}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow K(KK\phi \vee K\psi)\big) \overset{A_2,R_1,L.1}{\Longrightarrow} \big(K(K\phi \vee \psi) \rightarrow KK(K\phi \vee \psi)\big)$.

(A_4 for $\neg K\phi \vee \psi$) : The agent assumes A_5 for ϕ, implying $\big((\neg K\phi \vee \psi) \rightarrow (K\neg K\phi \vee \psi)\big) \overset{R_2}{\Longrightarrow} K\big((\neg K\phi \vee \psi) \rightarrow (K\neg K\phi \vee \psi)\big) \overset{A_2,R_1}{\Longrightarrow} \big(K(\neg K\phi \vee \psi) \rightarrow K(K\neg K\phi \vee \psi)\big) \overset{L.2,L.4,R_1}{\Longrightarrow} \big(K(\neg K\phi \vee \psi) \rightarrow (K\neg K\phi \vee K\psi)\big) \overset{L.2,R_1}{\Longrightarrow} K(\neg K\phi \vee \psi) \rightarrow (KK\neg K\phi \vee KK\psi) \overset{A_2,R_1,L.1}{\Longrightarrow} K(\neg K\phi \vee \psi) \rightarrow K(K\neg K\phi \vee K\psi) \overset{A_2,R_1,L.1}{\Longrightarrow} K(\neg K\phi \vee \psi) \rightarrow KK(\neg K\phi \vee \psi)$.

(A_5 for $K\phi \vee \psi$) : It follows from Lemma 4 that the agent has proven $(K\phi \vee K\psi) \rightarrow K(K\phi \vee \psi)$, implying $\big((K\phi \vee K\psi) \rightarrow K(K\phi \vee \psi)\big) \overset{A_1,R_1}{\Longrightarrow} \big(\neg K(K\phi \vee \psi) \rightarrow \neg(K\phi \vee K\psi)\big) \overset{A_1,R_1}{\Longrightarrow} \big(\neg K(K\phi \vee \psi) \rightarrow (\neg K\phi \wedge \neg K\psi)\big) \overset{A_5,R_1}{\Longrightarrow} \big(\neg K(K\phi \vee \psi) \rightarrow (K\neg K\phi \wedge K\neg K\psi)\big) \overset{L.1,R_1}{\Longrightarrow} \big(\neg K(K\phi \vee \psi) \rightarrow K(\neg K\phi \wedge \neg K\psi)\big) \overset{L.4,A_1,A_2,R_1}{\Longrightarrow} \big(\neg K(K\phi \vee \psi) \rightarrow K\neg K(K\phi \vee \psi)\big)$.

(A_5 for $\neg K\phi \vee \psi$) : It follows from Lemma 4 that the agent has proven $(K\neg K\phi \vee K\psi) \rightarrow K(K\neg K\phi \vee \psi)$. Hence, $\big((K\neg K\phi \vee K\psi) \rightarrow K(K\neg K\phi \vee \psi)\big) \overset{L.2,R_1}{\Longrightarrow} \big((K\neg K\phi \vee K\psi) \rightarrow K(\neg K\phi \vee \psi)\big) \overset{A_1,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow \neg(K\neg K\phi \vee K\psi)\big) \overset{A_1,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow (\neg K\neg K\phi \wedge \neg K\psi)\big) \overset{L.2,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow (K\neg K\neg K\phi \wedge K\neg K\psi)\big) \overset{L.1,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow K(\neg K\neg K\phi \wedge \neg K\psi)\big) \overset{A_1,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow K\neg(K\neg K\phi \vee K\psi)\big) \overset{L.4,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow K\neg K(K\neg K\phi \vee \psi)\big) \overset{L.2,R_1}{\Longrightarrow} \big(\neg K(\neg K\phi \vee \psi) \rightarrow K\neg K(\neg K\phi \vee \psi)\big)$. $\square$

Proof of Main Theorem. We recursively define $\Psi_n := \mathcal{L}^0(\{\phi, K\phi \mid \phi \in \Psi_{n-1}\})$ as the closure of $\{\phi, K\phi \mid \phi \in \Phi_{n-1}\}$ with respect to $\neg$ and $\wedge$, with $\Psi_0 := \mathcal{L}^0(\Phi)$ being the set of non-epistemic propositions. It follows directly from Lemmata 2, 3 and 5 that $A_3 - A_5$ for any proposition in Ψ_n are theorems in $S5(\Psi_{n-1})$, implying that the theorems in $S5(\Psi_n)$ and those in $S5(\Psi_{n-1})$ coincide. Therefore, every ML system $S5(\Psi_n)$ induces the same theorems as $S5^0$. Finally, notice that $\mathcal{L}(\Phi) = \bigcup_{n\geq 0} \Psi_n$, which completes the proof. □

3 Discussion

3.1 Partitional information structures

The standard modal logic system $S5$ is syntactic, in that it considers propositions and a knowledge operator. In order to make the connection between $S5$ (or $S5^0$) and partitional information, one needs to introduce a semantic model, given by states of the world.

The bridge between semantic and syntactic models consists of Kripke structures (Kripke, 1959), given as tuples $M = (\Omega, \pi, \mathcal{K})$; Ω is the set of states of nature; $\pi : \Omega \times \Phi \to \{0, 1\}$ is a function assigning a truth value to every primitive proposition, i.e., $\pi(\omega, p) = 1$ if and only if p is true at ω; $\mathcal{K} : \Omega \to 2^\Omega \setminus \{\emptyset\}$ determines a binary relationship on Ω, often called the agent's possibility correspondence, i.e., $\omega' \in \mathcal{K}(\omega)$ means that the agent deems the state ω' possible while being at ω. We write $(M, \omega) \vDash \phi$ whenever ϕ is true at ω in the Kripke structure M. Truth is defined inductively in M at every state as follows:

$(M, \omega) \vDash p$ for each $p \in \Phi$ if and only if $\pi(\omega, p) = 1$

$(M, \omega) \vDash \phi$ if and only if $(M, \omega) \nvDash \neg\phi$

$(M, \omega) \vDash \phi \wedge \psi$ if and only if $(M, \omega) \vDash \phi$ and $(M, \omega) \vDash \psi$

$(M, \omega) \vDash K\phi$ if and only if $(M, \omega') \vDash \phi$ for all $\omega' \in \mathcal{K}(\omega)$

Recall that M is reflexive whenever $\omega \in \mathcal{K}(\omega)$ for all $\omega \in \Omega$; it is transitive whenever for all $\omega, \omega' \in \Omega$, if $\omega' \in \mathcal{K}(\omega)$ then $\mathcal{K}(\omega') \subseteq \mathcal{K}(\omega)$; finally, it is Euclidean whenever for all $\omega, \omega' \in \Omega$, if $\omega' \in \mathcal{K}(\omega)$ then $\mathcal{K}(\omega') \supseteq \mathcal{K}(\omega)$. Partitional information structures correspond to Kripke structures that are reflexive, transitive and Euclidian.

A proposition ϕ is a tautology in M, and we write $M \vDash \phi$, whenever $(M, \omega) \vDash \phi$ for all $\omega \in \Omega$. It is valid in a class of Kripke structures $\mathcal{M}$, and we write $\mathcal{M} \vDash \phi$, whenever ϕ is a tautology in every $M \in \mathcal{M}$.

A modal logic system is a sound axiomatization of a class of Kripke structures $\mathcal{M}$ whenever every theorem in this ML system is a valid proposition in $\mathcal{M}$. A modal logic system is a complete axiomatization of $\mathcal{M}$ whenever every valid proposition in $\mathcal{M}$ is a theorem in the ML system. It is well known that $S5$ is a sound and complete axiomatization of partitional Kripke structures (Fagin et al., 1995, Ch. 3).

It immediately follows from our Main Theorem that $S5^0$ is also a sound and complete axiomatization of partitional Kripke structures.

3.2 Tightness of the result

We show that truth and introspection for the non-epistemic propositions are indispensable axioms, i.e., $A_3 - A_5$ cannot be proven only from $A_0 - A_2$, using semantic models as introduced in the previous subsection.

Suppose, for instance, that there is a unique atomic proposition, $\Phi = \{p\}$, and consider a Kripke structure M such that $\Omega = \{\omega, \omega'\}$, with $\pi(\omega, p) = 1$ and $\pi(\omega', p) = 0$. Moreover, let $\mathcal{K}(\omega) = \{\omega\}$ and $\mathcal{K}(\omega') = \{\omega, \omega'\}$. It is known that M belongs to the class of Kripke structures which are axiomatized by $A_1 - A_4$ together with the inference rules $R_1 - R_2$ (Fagin et al., 1995). Therefore, $\neg Kp \rightarrow K\neg Kp$ cannot be proven in M, implying that without assuming introspection for the non-epistemic propositions, we may obtain a strictly coarser set of theorems compared to $S5$.

Likewise, unless the truth axiom is assumed for the non-epistemic propositions, it cannot be proven by the remaining axioms. Consider, for instance, the following Kripke structure, M', such that $\Omega = \{\omega, \omega'\}$, with $\pi(\omega, p) = 1$ and $\pi(\omega', p) = 0$, and $\mathcal{K}(\omega) = \{\omega'\}$ and $\mathcal{K}(\omega') = \{\omega\}$. It is known that M' belongs to the class of Kripke structures which are axiomatized by $A_0 - A_2$, $A_4 - A_5$ together with the inference rules $R_1 - R_2$ (Fagin et al., 1995), implying that $\neg Kp \rightarrow p$ cannot be proven in M'.

3.3 Interactive knowledge

Let us extend our analysis to a framework with multiple agents $\{1, ..., n\}$, with typical elements i and j. Consider a separate knowledge modality K_i for every $i \in \{1, ..., n\}$, implying that the set of all propositions, denoted by $\mathcal{L}_n(\Phi)$, now becomes the closure of Φ with respect to $\neg$, $\wedge$ and $K_1, ..., K_n$, i.e., the language is enriched in order to contain propositions of the form "j knows that i knows p".

The ML system, $S5_n$, extends $S5$ to a multi-agent environment in a way such that all agents assume the basic axioms, e.g., $K_i\phi \rightarrow \phi$ is assumed for every $i \in \{1, ..., n\}$. The set of theorems[1] in $S5_n$ is denoted by T_{5n}.

Recall the definition of the non-epistemic propositions in the single-agent framework: It is the set of sentences that do not contain any knowledge operator. Extending this definition to the multi-agent framework should be done with caution, e.g., consider the proposition "j knows that i knows p": From j's point of view, K_jK_ip is epistemic, as it describes j's knowledge; on the other hand, K_jK_ip is a non-epistemic proposition in i's language as it describes j's mental state, even though the latter refers to i's knowledge. Hence, from i's point of view any proposition starting with K_j is non-epistemic.

Formally, we define the set of non-epistemic propositions in i's language

$$\mathcal{L}_i^0(\Phi) := \mathcal{L}^0\Big(\bigcup_{j\neq i} \{\ K_j\phi \mid \phi \in \mathcal{L}(\Phi)\ \} \cup \Phi\Big)$$

as the closure of $\bigcup_{j\neq i}\{K_j\phi \mid \phi \in \mathcal{L}(\Phi)\} \cup \Phi$ with respect to $\neg$ and $\wedge$. Obviously, when i is the only agent, it follows that $\mathcal{L}_i^0(\Phi) = \mathcal{L}^0(\Phi)$, implying that our generalized definition of non-epistemic propositions in the multi-agent environment is consistent with the single-agent case presented in the previous section.

Let $S5_n^0$ be the multi-agent generalization of $S5^0$:

A_1. All tautologies of propositional calculus

A_2. $\big(K_i\phi \wedge K_i(\phi \rightarrow \psi)\big) \rightarrow K_i\psi$

A_3. $K_i\phi \rightarrow \phi$, for all $\phi \in \mathcal{L}_i^0(\Phi)$

A_4. $K_i\phi \rightarrow K_iK_i\phi$, for all $\phi \in \mathcal{L}_i^0(\Phi)$

A_5. $\neg K_i\phi \rightarrow K_i\neg K_i\phi$, for all $\phi \in \mathcal{L}_i^0(\Phi)$

R_1. From ϕ and $(\phi \rightarrow \psi)$ infer ψ

R_2. From ϕ infer $K_i\phi$

Let T_{5n}^0 denote the theorems in $S5_n^0$.

Proposition 1. $T_{5n} = T_{5n}^0$.

[1] It is known that $S5_n$ is a sound and complete axiomatization of the class of multi-agent partitional Kripke structures $(\Omega, \pi, \mathcal{K}_1, ..., \mathcal{K}_n)$ (Fagin et al., 1995, Ch. 3).

Proof. Observe that $\mathcal{L}(\Phi)$ coincides with the closure of $\mathcal{L}^0_i(\Phi)$ with respect to $\neg$, $\wedge$ and K_i. Moreover, similarly to the Main Theorem, we show that i can prove $A_3 - A_5$ for all propositions in the closure of $\mathcal{L}^0_i(\Phi)$ with respect to $\neg$, $\wedge$ and K_i, and therefore i can prove $A_3 - A_5$ for all propositions in $\mathcal{L}(\Phi)$. Likewise, for every individual which completes the proof. □

It follows directly, from the previous result, that $S5^0_n$ is a sound and complete axiomatization of the class of multi-agent partitional Kripke structures.

Notice that in order to prove $A_3 - A_5$ for all propositions it does not suffice to assume the truth axiom and introspection only for $\mathcal{L}^0(\Phi)$, e.g., even if j assumes $(K_i\phi \to \phi)$, he cannot prove $(K_jK_i\phi \to K_i\phi)$. The reason is that, from j's point of view, $K_i\phi$ is a non-epistemic proposition, and therefore j cannot infer the truth axiom for $K_i\phi$.

References

Aumann, R.J. (1976). Agreeing to disagree. *Annals of Statistics* 4, 1236–1239.

——— (1999). Interactive epistemology I: Knowledge. *International Journal of Game Theory* 28, 263–300.

Chellas, B.F. (1980). *Modal logic: an introduction*. Cambridge University Press, Cambridge.

Fagin, R., Halpern, J.Y., Moses, Y. & Vardi, M.Y. (1995). *Reasoning about knowledge*. MIT press.

Geanakoplos, J. (1989). Game theory without partitions, and applications to speculation and consensus. *Cowles Foundation Discussion Paper* 914.

Halpern, J.Y.(1995). The effect of bounding the number of primitive propositions and the depth of nesting on the complexity of modal logic, *Artificial Intelligence* 75, 361–372.

Hintikka, J. (1962). *Knowledge and belief*. Cornell University Press, Ithaca, NY.

Kripke (1959). A completeness theorem in modal logic. *Journal of Symbolic Logic* 24, 1–14.

Samet, D. (1990). Ignoring ignorance and agreeing to disagree. *Journal of Economic Theory* 52, 190–207.