Introduction to Formal Epistemology Lecture 2a

Eric Pacuit and Rohit Parikh

August 14, 2007

Eric Pacuit and Rohit Parikh: Introduction to Formal Epistemology, Lecture 2a

イロト 不得下 イヨト イヨト ニヨー

\checkmark Introduction, Motivation and Basic Epistemic Logic

- Lecture 2: Other models of Knowledge, Knowledge in Groups and Group Knowledge
- Lecture 3: Reasoning about Knowledge and
- Lecture 4: Logical Omniscience and Other Problems
- Lecture 5: Reasoning about Knowledge in the Context of Social Software

▲ロト ▲暦 ト ▲ 臣 ト ▲ 臣 - ろんで

Epistemic Logic

The Language: $\varphi := p \mid \neg \varphi \mid \varphi \land \psi \mid K \varphi$

Kripke Models: $\mathcal{M} = \langle W, R, V \rangle$ with R an equivalence relation

Truth: $\mathcal{M}, w \models \varphi$ is defined as follows:

▲ロト ▲母 ト ▲目 ト ▲目 ト ▲ ● ろく⊙

Epistemic Logic

The Language: $\varphi := p \mid \neg \varphi \mid \varphi \land \psi \mid K \varphi$

Kripke Models: $\mathcal{M} = \langle W, R, V \rangle$ with R an equivalence relation

Truth: $\mathcal{M}, w \models \varphi$ is defined as follows:

▲ロト ▲母 ト ▲目 ト ▲目 ト ▲ ● ろく⊙

Epistemic Logic

The Language: $\varphi := p \mid \neg \varphi \mid \varphi \land \psi \mid K_i \varphi$ with $i \in \mathcal{A}$

Kripke Models: $\mathcal{M} = \langle W, \{R_i\}_{i \in \mathcal{A}}, V \rangle$ with R_i and equivalence relation.

Truth: $\mathcal{M}, w \models \varphi$ is defined as follows:

▲ロト ▲母 ト ▲目 ト ▲目 ト ● ● ● ● ● ●

Results

Modal Formula	Property	Philosophical Assumption
$K(\varphi ightarrow \psi) ightarrow (K \varphi ightarrow K \psi)$		Logical Omniscience
${\it K}arphi ightarrow arphi$	Reflexive	Truth
${\it K}arphi ightarrow{\it K}{\it K}arphi$	Transitive	Positive Introspection
eg K arphi ightarrow K eg K arphi K arphi	Euclidean	Negative Introspection
$ eg \kappa \bot$	Serial	Consistency

Results

The logic S5 contains the following axioms and rules:

$$\begin{array}{lll} Pc & \text{Axiomatization of Propositional Calculus} \\ K & K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi) \\ T & K\varphi \rightarrow \varphi \\ 4 & K\varphi \rightarrow KK\varphi \\ 5 & \neg K\varphi \rightarrow K\neg K\varphi \\ MP & \frac{\varphi & \varphi \rightarrow \psi}{\psi} \\ Nec & \frac{\varphi}{K\psi} \end{array}$$

Theorem

S5 is sound and strongly complete with respect to the class of Kripke frames with equivalence relations.

▲ロト ▲御 ト ▲ 臣 ト ▲ 臣 ト ● ○ ○ ○ ○ ○

Other Models

Aumann Structures

Group Knowledge

R. Aumann. Interactive Epistemology I & II. 1999.

▶ Let *W* be a set of worlds, or states.

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

- ▶ Let *W* be a set of worlds, or states.
- Let *S* be the set of all states of nature:

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

- Let W be a set of worlds, or states.
- Let S be the set of all states of nature: A state of nature is a complete description of the *exogenous* parameters (i.e. facts about the physical world) that do not depend on the agents' uncertainties.

▲ロト ▲暦 ト ▲ 臣 ト ▲ 臣 - ろんで

- ▶ Let *W* be a set of worlds, or states.
- Let *S* be the set of all states of nature.
- A set $E \subseteq W$, called an **event**, is true at state w if $w \in E$.

▲ロト ▲圖ト ▲画ト ▲画ト ▲画 - のへで

Definition

Aumann Model An **Aumann model based on S** is a triple $\langle W, \Pi, \sigma \rangle$, where W is a nonempty set, Π is a partition over W and $\sigma : W \to S$.

Definition

Knowledge Function Let $\mathcal{M} = \langle W, \Pi, \sigma \rangle$ be an Aumann model. The **knowledge function**, $K : \wp(W) \to \wp(W)$, based on \mathcal{M} is defined as follows:

$$\mathsf{K}(E) = \{w \mid \Pi(w) \subseteq E\}$$

▲ロト ▲御 ト ▲ 臣 ト ▲ 臣 ト ● ○ ○ ○ ○ ○

Lemma

Let $\mathcal{M} = \langle W, \Pi, \sigma \rangle$ be a Aumann model and K the knowledge function based on \mathcal{M} . For each $E, F \subseteq W$

$$E \subseteq F \Rightarrow K(E) \subseteq K(F) \qquad Mo$$

$$K(E \cap F) = K(E) \cap K(F) \qquad Clo
$$K(E) \subseteq E \qquad Tra
$$\frac{K(E) \subseteq K(K(E))}{K(E) \subseteq K(K(E))} \qquad Po$$

$$K(\emptyset) = \emptyset \qquad Co$$$$$$

Monotonicity Closure Under Intersection Truth Positive introspection Negative introspection Consistency

where \overline{E} means the set-theoretic complement of E (relative to W).

We can give analogous correspondence, completeness, etc. proofs.

▲口▶ ▲圖▶ ▲温▶ ▲温▶ 三連一

Bayesian Structures

- Let W be a set of worlds and ∆(W) be the set of probability distributions over W.
- We are interested in functions $p: W \to \Delta(W)$.
- The basic intuition is that for each state w ∈ W, p(w) ∈ Δ(W) is a probability function over W.
- So, p(w)(v) is the probability the agent assigns to state v in state w. To ease notation we write p_w for p(w).

◆□ ▶ ◆□ ▶ ◆□ ▶ ◆□ ▶ ◆□ ● ● ●

Definition

The pair $\langle W, p \rangle$ is called a **Bayesian frame**, where $W \neq \emptyset$ is any set, and $p: W \to \Delta(W)$ is a function such that

if
$$p_w(v) > 0$$
 then $p_w = p_v$

Given a Bayesian frame $\mathcal{F} = \langle W, p \rangle$ and a set of states *S*, an **Bayesian model based on S** is a triple $\langle W, p, \sigma \rangle$, where $\sigma : W \to S$.

Definition For each $r \in [0, 1]$ define $B^r : 2^W \to 2^W$ as follows

$$B^r(E) = \{w \mid p_w(E) \ge r\}$$

▲ロト ▲掛ト ▲ヨト ▲ヨト 三ヨ - のへで

Observation: We can define a possibility model from a Bayesian model as follows. Let $\langle W, p, \sigma \rangle$ be a Bayesian model on a state space *S*. We define a possibility model $\langle W, P, \sigma \rangle$ base on *S* as follows: define $\mathcal{P}: W \to 2^W$ by

$$\mathcal{P}(w) = \{v \mid \pi_w(v) > 0\}$$

It is easy to see that \mathcal{P} is serial, transitive and Euclidean.

◆□ ▶ ◆□ ▶ ◆□ ▶ ◆□ ▶ ◆□ ● ● ●

Other Models

Aumann Structures

Group Knowledge

Suppose there are two friends Ann and Bob are on a bus separated by a crowd. Before the bus comes to the next stop a mutual friend from outside the bus yells "get off at the next stop to get a drink?".

Say Ann is standing near the front door and Bob near the back door. When the bus comes to a stop, will they get off?

D. Lewis. Convention. 1969.

M. Chwe. Rational Ritual. 2001.

◆ロト ◆昼 ▶ ◆臣 ▶ ◆臣 ▶ ○ 臣 ○ のへ⊙

Suppose there are two friends Ann and Bob are on a bus separated by a crowd. Before the bus comes to the next stop a mutual friend from outside the bus yells "get off at the next stop to get a drink?".

Say Ann is standing near the front door and Bob near the back door. When the bus comes to a stop, will they get off?

D. Lewis. Convention. 1969.

M. Chwe. Rational Ritual. 2001.

Suppose there are two friends Ann and Bob are on a bus separated by a crowd. Before the bus comes to the next stop a mutual friend from outside the bus yells "get off at the next stop to get a drink?".

Say Ann is standing near the front door and Bob near the back door. When the bus comes to a stop, will they get off?

D. Lewis. Convention. 1969.

M. Chwe. Rational Ritual. 2001.

◆□ ▶ ◆□ ▶ ▲□ ▶ ▲□ ▶ ▲□ ● ● ●

Suppose there are two friends Ann and Bob are on a bus separated by a crowd. Before the bus comes to the next stop a mutual friend from outside the bus yells "get off at the next stop to get a drink?".

Say Ann is standing near the front door and Bob near the back door. When the bus comes to a stop, will they get off?

D. Lewis. Convention. 1969.

M. Chwe. Rational Ritual. 2001.

◆□ ▶ ◆□ ▶ ▲□ ▶ ▲□ ▶ ▲□ ● ● ●

Suppose there are two friends Ann and Bob are on a bus separated by a crowd. Before the bus comes to the next stop a mutual friend from outside the bus yells "get off at the next stop to get a drink?".

Say Ann is standing near the front door and Bob near the back door. When the bus comes to a stop, will they get off?

D. Lewis. Convention. 1969.

M. Chwe. Rational Ritual. 2001.

◆□ ▶ ◆□ ▶ ▲□ ▶ ▲□ ▶ ▲□ ● ● ●

Three Views of Common Knowledge

1. $\gamma := i$ knows that φ , j knows that φ , i knows that j knows that φ , j knows that i knows that φ , i knows that j knows that i knows that φ , ...

D. Lewis. Convention, A Philosophical Study. 1969.

2. $\gamma := i$ and j know that (φ and γ)

3. There is a shared situation s such that

- s entails φ
- s entails i knows arphi
- *s* entails *j* knows φ

H. Clark and C. Marshall. Definite Reference and Mutual Knowledge. 1981.

J. Barwise. Three views of Common Knowledge. TARK (1987)

◆ロト ◆昼 ▶ ◆臣 ▶ ◆臣 ▶ ○ 臣 ○ のへ⊙

Three Views of Common Knowledge

1. $\gamma := i$ knows that φ , j knows that φ , i knows that j knows that φ , j knows that i knows that φ , i knows that j knows that i knows that φ , ...

D. Lewis. Convention, A Philosophical Study. 1969.

- γ := i and j know that (φ and γ)
 G. Harman. *Review of* Linguistic Behavior. Language (1977).
- 3. There is a *shared situation s* such that
 - s entails φ
 - s entails i knows arphi
 - *s* entails *j* knows φ

H. Clark and C. Marshall. Definite Reference and Mutual Knowledge. 1981.

J. Barwise. Three views of Common Knowledge. TARK (1987)

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □

Three Views of Common Knowledge

1. $\gamma := i$ knows that φ , j knows that φ , i knows that j knows that φ , j knows that i knows that φ , i knows that j knows that j knows that φ , ...

D. Lewis. Convention, A Philosophical Study. 1969.

- γ := i and j know that (φ and γ)
 G. Harman. *Review of* Linguistic Behavior. Language (1977).
- 3. There is a *shared situation s* such that
 - s entails φ
 - s entails i knows φ
 - s entails j knows φ

H. Clark and C. Marshall. Definite Reference and Mutual Knowledge. 1981.

J. Barwise. Three views of Common Knowledge. TARK (1987).

・ロト ・ 御 ト ・ 注 ト ・ 注 ト

Recall $K_i : \wp(W) \to \wp(W)$ is *i*'s knowledge function.

Define
$$K^m : \wp(W) \to \wp(W)$$
 for $m \ge 1$ by
 $\mathsf{K}^1 E := \bigcap_{i \in \mathcal{A}} K_i E$
 $\mathsf{K}^{m+1} E := \mathsf{K}^1(\mathsf{K}^m(E))$

 K^1E means everyone knows E K^2E means everyone knows that everyone knows E

Define $K^{\infty} : \wp(W) \to \wp(W)$

$K^{\infty}E := K^1E \cap K^2E \cap \cdots \cap K^mE \cap \cdots$

◆ロト ◆昼 ▶ ◆臣 ▶ ◆臣 ▶ ○ 臣 ○ のへ⊙

Recall $K_i : \wp(W) \to \wp(W)$ is *i*'s knowledge function.

Define
$$K^m : \wp(W) \to \wp(W)$$
 for $m \ge 1$ by
 $\mathsf{K}^1 \mathsf{E} := \bigcap_{i \in \mathcal{A}} \mathsf{K}_i \mathsf{E}$
 $\mathsf{K}^{m+1} \mathsf{E} := \mathsf{K}^1(\mathsf{K}^m(\mathsf{E}))$

 K^1E means everyone knows E K^2E means everyone knows that everyone knows E

Define $K^{\infty} : \wp(W) \to \wp(W)$

$K^{\infty}E := K^1E \cap K^2E \cap \cdots \cap K^mE \cap \cdots$

◆□▶ ◆□▶ ★□▶ ★□▶ □ のQ@

Recall $K_i : \wp(W) \to \wp(W)$ is *i*'s knowledge function.

Define
$$K^m : \wp(W) \to \wp(W)$$
 for $m \ge 1$ by
 $\mathsf{K}^1 \mathsf{E} := \bigcap_{i \in \mathcal{A}} \mathsf{K}_i \mathsf{E}$
 $\mathsf{K}^{m+1} \mathsf{E} := \mathsf{K}^1(\mathsf{K}^m(\mathsf{E}))$

 K^1E means everyone knows E K^2E means everyone knows that everyone knows E

Define $K^{\infty} : \wp(W) \to \wp(W)$

$$K^{\infty}E := K^1E \cap K^2E \cap \cdots \cap K^mE \cap \cdots$$

▲ロト ▲御 ト ▲ 臣 ト ▲ 臣 ト ● のへで

Recall $K_i : \wp(W) \to \wp(W)$ is *i*'s knowledge function.

Define
$$K^m : \wp(W) \to \wp(W)$$
 for $m \ge 1$ by
 $\mathsf{K}^1 \mathsf{E} := \bigcap_{i \in \mathcal{A}} \mathsf{K}_i \mathsf{E}$
 $\mathsf{K}^{m+1} \mathsf{E} := \mathsf{K}^1(\mathsf{K}^m(\mathsf{E}))$

 K^1E means everyone knows E K^2E means everyone knows that everyone knows E

Define $K^{\infty}: \wp(W) \rightarrow \wp(W)$

$K^{\infty}E := K^1E \cap K^2E \cap \cdots \cap K^mE \cap \cdots$

◆□▶ ◆□▶ ★□▶ ★□▶ □ のQ@

Recall $K_i : \wp(W) \to \wp(W)$ is *i*'s knowledge function.

Define
$$K^m : \wp(W) \to \wp(W)$$
 for $m \ge 1$ by
 $\mathsf{K}^1 \mathsf{E} := \bigcap_{i \in \mathcal{A}} \mathsf{K}_i \mathsf{E}$
 $\mathsf{K}^{m+1} \mathsf{E} := \mathsf{K}^1(\mathsf{K}^m(\mathsf{E}))$

 K^1E means everyone knows E K^2E means everyone knows that everyone knows E

Define $K^\infty : \wp(W) \to \wp(W)$

$$K^{\infty}E := K^1E \cap K^2E \cap \cdots \cap K^mE \cap \cdots$$

◆□▶ ◆□▶ ★□▶ ★□▶ □ のQ@

Fact Prove that for all $i \in A$ and $E \subseteq W$, $K_i K^{\infty}(E) = K^{\infty}(E)$.

▲口 ▶ ▲ 圖 ▶ ▲ 国 ▶ ▲ 国 ▶ ▲ 国 ▶ ▲ 国 ▶

Fact Prove that for all $i \in A$ and $E \subseteq W$, $K_i K^{\infty}(E) = K^{\infty}(E)$.

Suppose you are told "Ann and Bob are going together,"' and respond "sure, that's common knowledge." What you mean is not only that everyone knows this, but also that the announcement is pointless, occasions no surprise, reveals nothing new; in effect, that the situation after the announcement does not differ from that before. ...the event "Ann and Bob are going together" — call it E — is common knowledge if and only if some event call it F — happened that entails E and also entails all players' knowing F (like all players met Ann and Bob at an intimate party). (Aumann, pg. 271, footnote 8)

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □

Definition

Self-Evident Event An event F is **self-evident** if $K_i(F) = F$ for all $i \in A$.

Definition

Knowledge Field Let $\langle W, \{\Pi_i\}_{i \in \mathcal{A}}, \sigma \rangle$ be a multi-agent Aumann model. For each $i \in \mathcal{A}$, the **knowledge field of** *i*, denoted \mathbb{K}_i , is the family of all unions of cells in Π_i .

Lemma

An event E is commonly known iff some self-evident event that entails E obtains. Formally, $K^{\infty}(E)$ is the largest event in $\bigcap_{i \in A} \mathbb{K}_i$ that is included in E.

◆□▶ ◆□▶ ★□▶ ★□▶ □ のQ@
Definition

The operator "everyone knows φ ", denoted $E\varphi$, is defined as follows

$$E\varphi := \bigwedge_{i\in\mathcal{A}} K_i \varphi$$

< 日 > (同 > (回 > (回 >))) (回 >) (回 >)) (回 >) (u =

Definition

The operator "everyone knows φ ", denoted $E\varphi$, is defined as follows

$$E arphi := \bigwedge_{i \in \mathcal{A}} K_i arphi$$

Definition

The multi-agent epistemic language with common knowledge is generated by the following grammar:

$$p \mid \neg \varphi \mid \varphi \land \varphi \mid K_i \varphi \mid C \varphi$$

where $p \in At$ and $i \in A$.

◆□ ▶ ◆□ ▶ ▲□ ▶ ▲□ ▶ ▲□ ● ● ●

Definition The truth of $C\varphi$ is:

 $\mathcal{M}, w \models C \varphi$ iff for all $v \in W$, if wR^*v then $\mathcal{M}, v \models \varphi$

where $R^* := (\bigcup_{i \in A} R_i)^*$ is the reflexive transitive closure of the union of the R_i 's.

▲ロト ▲掛ト ▲ヨト ▲ヨト 三ヨ - のへで

Definition The truth of $C\varphi$ is:

 $\mathcal{M}, w \models C \varphi$ iff for all $v \in W$, if wR^*v then $\mathcal{M}, v \models \varphi$

where $R^* := (\bigcup_{i \in A} R_i)^*$ is the reflexive transitive closure of the union of the R_i 's.

 $\mathcal{M}, w \models C\varphi$ iff every finite path starting at w ends with a state satisfying φ .

◆□▶ ◆□▶ ★□▶ ★□▶ □ のQ@

Example



P means "The talk is at 2PM".

Eric Pacuit and Rohit Parikh: Introduction to Formal Epistemology, Lecture 2a

▲□▶ ▲圖▶ ▲臣▶ ▲臣▶ 二臣…

Example



P means "The talk is at 2PM".

 $\mathcal{M}, s \models \neg K_B P$

Example



P means "The talk is at 2PM".

 $\mathcal{M}, s \models \neg K_B P$

Example



P means "The talk is at 2PM".

 $\mathcal{M}, s \models E \neg K_B P$

Example



P means "The talk is at 2PM".

 $\mathcal{M}, s \models E \neg K_B P$

Example



P means "The talk is at 2PM".

 $\mathcal{M}, s \models C \neg K_B P$

Two players Ann and Bob are told that the following will happen. Some positive integer n will be chosen and *one* of n, n + 1 will be written on Ann's forehead, the other on Bob's. Each will be able to see the other's forehead, but not his/her own.

Suppose the number are (2,3).

Do the agents know there numbers are less than 1000?

Is it common knowledge that their numbers are less than 1000?

Two players Ann and Bob are told that the following will happen. Some positive integer n will be chosen and *one* of n, n + 1 will be written on Ann's forehead, the other on Bob's. Each will be able to see the other's forehead, but not his/her own.

Suppose the number are (2,3).

Do the agents know there numbers are less than 1000?

Is it common knowledge that their numbers are less than 1000?

Two players Ann and Bob are told that the following will happen. Some positive integer n will be chosen and *one* of n, n + 1 will be written on Ann's forehead, the other on Bob's. Each will be able to see the other's forehead, but not his/her own.

Suppose the number are (2,3).

Do the agents know there numbers are less than 1000?

Is it common knowledge that their numbers are less than 1000?

Two players Ann and Bob are told that the following will happen. Some positive integer n will be chosen and *one* of n, n + 1 will be written on Ann's forehead, the other on Bob's. Each will be able to see the other's forehead, but not his/her own.

Suppose the number are (2,3).

Do the agents know there numbers are less than 1000?

Is it common knowledge that their numbers are less than 1000?

Logics with Common Knowledge

The following axiom and rule need to be added to $\mathbf{S5}$ to deal with the common knowledge operator:

$$\blacktriangleright \ C(\varphi \to \psi) \to (C\varphi \to C\psi)$$

$$\blacktriangleright \ C\varphi \to (\varphi \land EC\varphi)$$

$$\blacktriangleright \ C(\varphi \to E\varphi) \to (\varphi \to C\varphi)$$

Theorem

S5^C is sound and weakly complete with respect to the class of all Kripke frames where the relations are equivalence relations.

Logics with Common Knowledge

The following axiom and rule need to be added to ${f S5}$ to deal with the common knowledge operator:

$$\blacktriangleright \ C(\varphi \to \psi) \to (C\varphi \to C\psi)$$

$$\blacktriangleright \ C\varphi \to (\varphi \land EC\varphi)$$

$$\blacktriangleright \ C(\varphi \to E\varphi) \to (\varphi \to C\varphi)$$

Theorem

S5^{*C*} is sound and weakly complete with respect to the class of all Kripke frames where the relations are equivalence relations.

- In Relational Models (including Aumann Structures), the first two views of Common Knowledge are mathematically equivalent and it is not clear how to represent the third view.
- They can be separated using alternative semantics.
 J. van Benthem and D. Sarenac. The Geometry of Knowledge. 2005.
 J. Barwise. Three Views of Common Knowledge. TARK, 1987.
- There are other notions of "group knowledge": distributed knowledge
- In many social situations, other levels of knowledge are of interest...

- In Relational Models (including Aumann Structures), the first two views of Common Knowledge are mathematically equivalent and it is not clear how to represent the third view.
- They can be separated using alternative semantics.
 J. van Benthem and D. Sarenac. The Geometry of Knowledge. 2005.
 J. Barwise. Three Views of Common Knowledge. TARK, 1987.
- There are other notions of "group knowledge": distributed knowledge
- In many social situations, other levels of knowledge are of interest...

- In Relational Models (including Aumann Structures), the first two views of Common Knowledge are mathematically equivalent and it is not clear how to represent the third view.
- They can be separated using alternative semantics.
 J. van Benthem and D. Sarenac. The Geometry of Knowledge. 2005.
 J. Barwise. Three Views of Common Knowledge. TARK, 1987.
- There are other notions of "group knowledge": distributed knowledge
- In many social situations, other levels of knowledge are of interest...

- In Relational Models (including Aumann Structures), the first two views of Common Knowledge are mathematically equivalent and it is not clear how to represent the third view.
- They can be separated using alternative semantics.
 J. van Benthem and D. Sarenac. The Geometry of Knowledge. 2005.
 J. Barwise. Three Views of Common Knowledge. TARK, 1987.
- There are other notions of "group knowledge": distributed knowledge
- In many social situations, other levels of knowledge are of interest...

- In Relational Models (including Aumann Structures), the first two views of Common Knowledge are mathematically equivalent and it is not clear how to represent the third view.
- They can be separated using alternative semantics.
 J. van Benthem and D. Sarenac. The Geometry of Knowledge. 2005.
 J. Barwise. Three Views of Common Knowledge. TARK, 1987.
- There are other notions of "group knowledge": distributed knowledge
- In many social situations, other levels of knowledge are of interest...

◆□ ▶ ◆□ ▶ ▲□ ▶ ▲□ ▶ ▲□ ● ● ●

Time for a beak.

A travelling salesman found himself spending the night at home with his wife when one of his trips was accidentally cancelled. The two of them were sound asleep, when in the middle of the night there was a loud knock at the front door. The wife woke up with a start and cried out, "Oh my God! It's my husband!" Whereupon the husband leapt out of bed, ran across the room and jumped out the window.

Schank and Abelson, 1977, p. 59.

Wimmer and Perner on Beliefs about beliefs

Mother returns from her shopping trip. She bought chocolate for a cake. Maxi may help her put away the things. He asks her, "Where should I put the chocolate?" "In the blue cupboard," says the mother.

Later, with Maxi gone out to play, the mother transfers the chocolate from the blue cupboard to the green cupboard. Maxi then comes back from the playground, hungry, and he wants to get some chocolate.

In Wimmer and Perner's experiment, little children who were told the Maxi story were then asked the *belief* question, "Where will Maxi look for the chocolate?" Children at the age of three or less invariably got the answer wrong and assumed that Maxi would look for the chocolate in the *green* cupboard where *they* knew it was. Even children aged four or five had only a onethird chance of correctly answering this question or an analgous question involving Maxi and his brother (who also wants the chocolate and whom Maxi wants to deceive).

Children aged six or more were by contrast quite successful in realizing that Maxi would think the chocolate would be in the *blue* cupboard – where he had put it and that if he wanted to deceive his brother, he would lead him towards the green cupboard.

The Muddy Children Puzzle

In this by now well-known puzzle, a number of children are playing in the mud and some of them get their foreheads dirty. At this the father comes on the scene and announces, "at least one of you has got her forehead dirty."

Scenario 1: Suppose there is only one child, say Amy, who is dirty. Then she will realize that her own forehead must be dirty since she can see that the others are clean.

Scenario 2: Suppose now that there are two dirty children, Sarah and Amy, who are asked in turn, "Do you know if your forehead is dirty?" Now when Sarah is asked, she can see Amy's dirty forehead and she replies, "I don't know." However, when Amy is asked, she is able to reason, "If my forehead were clean, Sarah would have known that hers must be dirty since all the others are clean. But Sarah did not know. So my forehead must be dirty."

This reasoning on Amy's part requires a representation by Amy of Sarah's state of mind, and clearly Amy must be at least six for this to work. However, Sarah herself must have some reasoning ability and Amy must know that she has such abilities. It is not enough for Amy to know Sarah's view of reality, she must also represent Sarah's logical abilities in her own mind.

In particular, suppose that there are three dirty children – Jennifer, Sarah, and Amy – who are asked in turn whether they know if they are dirty, and with Amy being asked last. If Sarah is only three, Amy would not be justified in concluding from Sarah's "I don't know" that in that case, Amy herself must be dirty. Amy would need to know that *if* Amy were clean, Sarah would have carried out a representation in her own mind of Jennifer's state of mind and concluded from Jennifer's "I don't know" that Sarah must herself be dirty. But if Sarah is only three, Amy cannot rely on such reasoning on Sarah's part.

Man or Elephant?

This example comes from the *Mahabharata*, one of the Indian epics and at 100,000 verses, believed to be the longest single work in the ancient world. It describes the political struggle between two sets of cousins, the *Pandavas*, and the *Kauravas*. In a crucial battle between the two sets of cousins, *Krishna* is an adviser to the *Pandavas*, but out of a sense of fair play he gives his army to the *Kauravas*. At a crucial juncture *Drona*, a powerful warrior on the *Kauravas* and *Kauravas*, turns out to be invincible in battle and the *Pandavas* are hard pressed.

However, the wily Krishna thinks up a strategem.

Drona's only son is called Ashvatthama and so is an elephant owned by the Pandavas. Krishna suggests that the Pandavas kill the elephant Ashvatthama and then announce that the man Ashvatthama has been killed.

After a great deal of hesitation and soul searching, selfinterest prevails, the *Pandavas* do kill the elephant, and announce to *Drona* that *Ashvatthama* is dead. Drona is a little suspicious, but knows that one of the *Pandava* brothers, *Yudhisthira* never lies. He asks *Yudhisthira*, who confirms that *Ashvatthama* is dead, muttering in an aside, "either man or elephant." Not knowing about the elephant, *Drona* assumes it is his son who is dead, lays down his weapons, and is killed by a warrior on the *Pandava* side. We now offer a game-knowledge theoretic analysis of this event.

Let *Drona's* two options be f for 'fight' and n for not fight. Before the announcement his preferences were f > n and given his prowess, any warrior who faced him faced death.

But after the announcement that Ashwatthama was dead, his preferences change to n > f, and he can be attacked with impunity.

Ahswatthama of course was not dead and took terrible revenge on the *Pandavas* including the killing of some unborn *Pandava* children. For this he was punished by being condemned to live forever, and to wander the earth as a pariah.

The Pedestrian and the Motorist

Deception consists of inducing a second-order false belief.

Let S be the situation where a pedestrian is crossing the street and a car is coming.

Let S' be the same situation without the car.

		G	Ν
Pedestrian choices	g	(-100,-10)	(1,0)
	n	(0,1)	(0,0)

Motorist choices

Figure I

There are two Nash equilibria: at (g, N) and at (n, G). However, the penalty for the pedestrian (injury or loss of life) to depart from (n, G) is much greater than the penalty for the motorist (fine or loss of license) to depart from (g, N). Thus the equilibrium (g, N) is less stable than (n, G), and this fact creates the possibility for the motorist to 'bully' the pedestrian. However, if the pedestrian is unaware of the existence of the car, then the picture is much simpler and his payoffs are 1 for g and 0 for n. g dominates n, and once this choice is made by the pedestrian, it is dominant for the motorist to choose N. This is why the pedestrian tries to achieve the state of knowledge represented by the formulas $K_p(C)$, $\neg K_m(K_p(C))$ indicating that the pedestrian knows the car is there but the motorist does not know that the pedestrian knows. The pedestrian chooses the action g, and knowing that the pedestrian will do this, the motorist must choose N. However, if the motorist has a horn, he can change the knowledge situation. The existence of the car becomes common knowledge and thus the possibility for the motorist to bully the pedestrian arises again.

The Problem of the Two Generals (Halpern and Moses)

In this problem there are two generals A, B who are stationed on opposing hilltops and who wish to attack an enemy E in the valley below. General A sends a messenger to general B suggesting that the two should attack together at dawn. However, there is a possibility that the message might not reach B. Perhaps the messenger would be captured or killed by E, and general A attacking by himself would be badly defeated. So general A asks for an acknowledgement from B that his message *let us attack at dawn* has been received.

However, general B has the same problem. He does not wish to attack alone and so he agrees to the attack, but asks for an acknowledgement in turn. Clearly this process of saying *attack at dawn, please acknowledge* can never end. What is needed, Halpern and Moses argue, is common knowledge of the intended attack and no finite number of messages back and forth will achieve it. However, there is a small twist in this problem to which we now turn.

Let a stand for the event that general A attacks alone. Let b stand for the event that general B attacks alone. Let t stand for the event that they attack together. And finally let n stand for the event that neither attacks.

The analysis which Halpern and Moses provide tacitly assumes that the priorities are: t > n > b > a for general A, and t > n > a > b for general B. "If only one general is going to attack, let it be the other guy," is tacitly assumed.

Suppose, however, that t > b > n > a for both generals. Then general A can issue the message "Let us plan to attack at dawn, please acknowledge, but I will not acknowledge your acknowledgement."

In such a case, A is guaranteed not to attack alone, and there may well be a good chance of t, provided only that the probability of a message getting through is high enough. As our final example we consider the ballot box whose function is to create certain specific states of knowledge. Suppose that five people 1, 2, 3, 4, 5 are debating whether to have lettuce or cucumbers for salad. They cast their votes into a ballot box and the final count reveals that lettuce has won by 3 votes to 2.

Let $C_i : i \leq 5$ mean that *i* voted for cucumber, and similarly for lettuce. Then the propositional formula which expresses that exactly two voted for lettuce and three for cucumbers is common knowledge. This is a formula with 10 disjunctions, a typical one being $(L_1 \wedge L_2 \wedge C_3 \wedge L_4 \wedge C_5)$. Also what is common knowledge is $(\forall i, \forall j) (i \neq j \rightarrow \neg K_i(C_j))$ as well as $(\forall i, \forall j) (i \neq j \rightarrow \neg K_i(L_j))$. I.e. it is common knowledge that no one knows anyone else's vote. I am sure the reader can see the game theoretic reasons for these two knowledge facts. Everyone must know the results of the election, and the lettuce party must not be in a position to take revenge on the two cucumbers.

Common Knowledge and Consensus

In 1976 Aumann proved the following remarkable result. Suppose two agents start out with a common partition of the state space W but then receive different information. Let A be an event, and let p, q be the new probabilities of A for the two agents.

Theorem 1: If p and q are common knowledge, then p = q.

Aumann's result used the sure thing principle. Let E, F be disjoint events and suppose that p(A|E) = p(A|F) = x. Then $p(A|E \cup F) = x$.

Geanakoplos and Polemarchakis showed that if p and q were not equal, but the agents exchanged their values of p and q, updating properly, then their values of p and q updated would become equal after a finite number of steps.

Corollary (to Aumann): It is not possible for two agents Ann and Bob to be such that Ann thinks some stock is going down and Bob thinks that the stock is going up, so that Ann is selling the stock to Bob.
Definition 1: Let f be a function defined on subsets of W into the reals. f is *convex* if a) f satisfies the sure thing principle and b) if f(E) < f(F) then $f(E) < f(E \cup F) < f(F)$.

Theorem 2: (RP and Krasucki) If n agents exchange values of such a convex f in pairwise communications so that no agent is left out of the chain of communication, then eventually their estimates of value of f will become equal.

Levels of Knowledge

Given a group $N = \{1, ..., n\}$ of agents (whether people or processes), what are the properties of their state of knowledge relative to some fact A?

Definition 2: Assume given m propositional variables $P_1, ..., P_m$. Let $L_0 = \{P_1, ..., P_m\}$ and let L_g be all boolean combinations of the P_i . If the P_i are *basic* ground facts, then L_g represents all ground (knowledge-free) facts. Given a group $N = \{1, ..., n\}$ of agents, we define the full knowledge language L as follows:

- (i) $L_0 \subseteq L$
- (ii) If $A, B \in L$ then so are $\neg A, A \lor B$.
- (iii) If $A \in L$ then for all $i \leq n, K_i(A) \in L$.

To consider common knowledge as well we extend L to L_c by adding the conditions (iv) $L \subseteq L_c$ (v) If $A \in L_c$ and $U \subseteq N$ then $C_U(A) \in L_c$. For convenience we shall identify K_i with $C_{\{i\}}$. **Definition 3:** A Kripke structure \mathcal{M} for L consists of a nonempty set W of states, a map π from $W \times L_0$ into $\{1,0\}$ with 1 standing for true and 0 for false, and finally an equivalence relation R_i over W for each $i \leq m$.

Definition 4: Given a Kripke structure \mathcal{M} for L, a state $s \in W$ and a formula $A \in L_c$ we define $\mathcal{M}, s \models A$ as follows by induction on the complexity of A. First, for each $U \subseteq N$, we define the relation R_U to be the transitive closure of $\bigcup R_i : i \in U$. Then we have: (i) If A is atomic then $\mathcal{M}, s \models A$ iff $\pi(s, A) = 1$ (ii) If $A = \neg B$ then $\mathcal{M}, s \models A$ iff $\mathcal{M}, s \not\models B$ (iii) If $A = B \lor C$ then $\mathcal{M}, s \models A$ iff $\mathcal{M}, s \models B$ or $\mathcal{M}, s \models C$ (iv) If $A = C_a(B)$ where a is either some i or else some U, then $\mathcal{M}, s \models A$ iff $(\forall t)((s, t) \in R_a \to \mathcal{M}, t \models B)$ **Theorem 3:** Let Σ_C be the alphabet whose symbols are $\{C_U\}_{U\subseteq N}$ For all x, y in Σ_C^* , and all formulae A, for all \mathcal{M}, s , $V \subseteq U \subseteq N$, $\mathcal{M}, s \models xC_UC_VyA$ iff $\mathcal{M}, s \models xC_VC_UyA$ iff $\mathcal{M}, s \models xC_UyA$.

In other words, common knowledge by the larger group U absorbs common knowledge by the smaller one.

Corollary 1: Let Σ_K be the alphabet whose symbols are $\{K_1, \ldots, K_n\}$. For all $a = K_i$ in Σ_K , and for all x, y, in Σ_K^* , and all formulae A,

$$\models xayA \leftrightarrow xaayA$$

and hence for all $\mathcal{M}, s, \mathcal{M}, s \models xayA$ iff $\mathcal{M}, s \models xaayA$, i.e., repeated occurrences of a are without effect and if $xay \in L_K(A, s)$ then $\forall n \ xa^n y \in L_K(A, s)$.

In other words, it is common knowledge that a knowing some B is the same as a knowing that a knows B.

Definition 5: Given a formula A and \mathcal{M}, s the *level* of A at s, L(A, s) is the set of x in Σ_C^* such that $\mathcal{M}, s \models xA$, and x contains no substrings $C_U C_V, C_V C_U$ for any $V \subseteq U \subseteq N$.

Strings x such that x contains no substrings $C_U C_V$, $C_V C_U$ for any $V \subseteq U \subseteq N$ will be called *simple*.

Embeddability

Definition 6: Given two strings $x, y \in \Sigma_K^*$, we say that x is *embeddable* in y ($x \leq y$), if all the symbols of x occur in y, in the same order, but not necessarily consecutively. Formally:

1) $x \le x, \epsilon \le x$ for all x2) $x \le y$ if there exist x', x'', y', y'', such that x = x'x'', y = y'y'', and $x' \le y', x'' \le y''$.

and \leq is the smallest relation satisfying (1) and (2).

Thus the string aba is embeddable in itself, in aaba and in abca, but not in aabb.

Fact 1: Embeddability is a *well-partial order*, i.e. it is not only well-founded, but every linear order that extends it is a well-order. Equivalently, it is well-founded and every set of mutually incomparable elements is finite.

Note for instance that an infinite set of incomparable elements $\{a_1, ..., a_n, ...\}$ is well-founded – nothing is below anything else. However, it is not a WPO, for we can clearly set $a_1 > a_2 > a_3...$ which gives an extension of the original, flat ordering.

Fact 2: Embeddability can be tested in linear time, e.g., by a nondeterministic finite automaton with two input tapes.

Fact 1 was proved first by Graham Higman [Hi]. See [JP] for a discussion. Fact 2 is straightforward.

We also need a stronger relation defined on Σ_C^* , which we call *C*-embeddability.

Definition 7: Given two strings $x, y \in \Sigma_C^*$, we say that x is C-embeddable in $y \ (x \leq y)$, if 1) If $V \subseteq U$ then $C_V \leq C_U$ 2) $x \leq y$ if there exist $x', x'', y', y'', (y', y'' \neq \epsilon)$, such that x = x'x'', y = y'y'', and $x' \leq y', x'' \leq y''$. and \leq is the smallest relation satisfying (1) and (2).

Fact 3: For any $x, y \in \Sigma_K^*$, $x \leq y$ iff $x \leq y$.

Fact 4: C-embeddability is a well-partial order.

Definition 8: $R \subseteq S$ is downward closed iff $x \in R$ implies $\forall y \leq x, y \in R$.

The Main Results on Levels of Knowledge

Theorem 4: Let Σ_C be the alphabet whose symbols are $\{C_U\}_{U\subseteq N}$. Then for all strings x, y in Σ_C^* , if $x \leq y$ then for all \mathcal{M}, s , if $\mathcal{M}, s \models yA$ then $\mathcal{M}, s \models xA$.

Corollary 1: Every level of knowledge is a downward closed set with respect to \leq .

Theorem 5: There are only countably many levels of knowledge and in fact all of them are regular subsets of Σ^* (where Σ is either Σ_K or Σ_C).

Fact 5: Eric Pacuit and ourselves have shown that in contrast with *knowledge* there are *uncountably* many possible levels of rational *belief*. This is curious, as truth is the only condition which (formally) separates knowledge from rational belief.

Corollary 1: The membership problem for a level of knowledge can be solved in linear time.

Now we consider what finite downward closed sets of strings can look like.

Theorem 6: If L is a non-empty finite subset of Σ_K^* , then L is downward closed iff for some k,

$$L = \bigcup_{i=1}^{k} dc(\{x_i\})$$

where $x_i \in \Sigma_K^*$.

Proof: Consider the set M of maximal elements of L. Then because the order is a WPO, the set M must be finite. Moreover, every element of L must lie below some maximal element. Hence if $M = \{x_1, ..., x_k\}$ then we get $L = \bigcup_{i=1}^k dc(\{x_i\})$.

This theorem reiterates the fact that the finite levels are characterized by their maximal elements $(x_1, ..., x_k$ are maximal). The characterization of infinite levels of knowledge is more complex. The details are in [PK]. We have shown that every level of knowledge is a regular set of strings satisfying certain conditions. But do all such sets actually arise as levels of knowledge? We now give a simple argument to show that they do in fact. The following result is proved jointly with Eric Pacuit.

Theorem 7: Let L be a downward closed set of strings relative to \leq . Then there is a finite Kripke model Mand state s such that for all strings x in $\{K_1, ..., K_n\}^*$, $M, s \models xp$ iff $x \in L$, where p is a propositional variable.

Applications to Games

Game G: This is really a pair of games, but having warned the reader we will just call it a game. In this game each player chooses a number between 1 and 10. If the two numbers are more than 1 apart, the payoff is 0 in both M, N. If the numbers chosen are at most 1 apart, then in N the payoff to each player is the *maximum* of the two numbers, say a, b. However, if they are at most 1 apart, then the payoff in M will be 10 - min(a, b).

So in N it pays to pick the higher numbers, and in M it pays to pick the lower numbers. Thus for instance the numbers (2,6) will yield 0 payoff in both matrices, whereas (2,3) will yield a payoff of 3 in N and of 8 in M.

We assume that N is the *default game* but that the game *could* be M. p is the proposition that the actual game is M.

Now if C does not know p, C – assuming that the matrix is N – will play 10. In N, this would give the highest possible payoff of 10, provided that R plays 9 or 10. Now if R knows p and that C does not know p, R will know that C will play 10 and R herself will play 9. Thus the payoff to each will be 10 - 9 = 1.

Suppose now that C knows that R knows p, but R does not know this. Then C will know that R will play 9 and C will play 8, with the payoff to both being 2. R will be surprised, but pleasantly.

As the level of knowledge of p goes up, so will the payoffs, until a maximum of 10 is reached with one of the players playing 0 and the other playing 1. Even though the payoffs are co-ordinated, higher levels of knowledge bring greater benefits, until level 10 is reached and after that, even common knowledge will be no better.

In any case, it is evident that in finite strategy games with co-ordinated payoffs, the higher levels of knowledge always bring better payoffs. This will be the case in two player games of co-ordination, provided only that one player knows everything which the other does. The situation is different if there are more than two players or if there are infinitely many strategies for both players. With more than two players a problem can arise if players have incomparable knowledge. Thus if there are three players A, B, C, and A and B have different notions of what C will play, then even though the three have common interests, A and B might make choices which will make the outcome worse for all three. But if there is a hierarchy so that player A knows everything which B knows, and player B knows everything which C knows, and the games are co-ordinated so that a benefit to one is a benefit to the others, then we will still have the result that higher, *finite* levels are beneficial. **Game G2:** In this game both players play some natural number. The payoff is 0 in both M, N if the difference between the two numbers is more than 1. In N, if either number exceeds 10, the payoff is 0, but if both numbers are ≤ 10 and no more than 1 apart, the payoff is the maximum of the two numbers. Matrix M is similar to matrix N in that higher numbers are better, but there is no punishment for numbers > 10. The payoff is the maximum, period, provided only that the numbers are no more than 1 apart.

Suppose now that p is true, R knows it but C does not. Then C will play 10 and knowing this, R will play 11. The payoff will be 11 for both.

If $K_r(p)$, $K_c(K_r(p))$, $K_c(\neg K_r(K_c(p)))$, then C will know that R will play 11 and C himself will play 12, thus getting a payoff of 12 for both. As the level of knowledge rises, so will the payoff. But now there is a paradox! If p is common knowledge, the players will have no idea how to play! So common knowledge is not necessarily better than a high finite level of knowledge. **Non-coordinated games:** Suppose now that we are dealing with games where the payoffs are not positively correlated. Perhaps one or both matrices are zero sum, although as we noticed, it is really the matrix M which counts. Matrix N is only used to establish that 1 is the default strategy for C.

Now suppose we have some level of knowledge $K_r(p)$, $K_c(K_r(p))$, rising to some finite level. Then we may have the situation that the default play for C (in N) is 1, the best response to that (in M, the actual game) is 2, the best response to that is 3 and so on, until we reach the play corresponding to the actual level of knowledge, perhaps (r(2n, 2n + 1), c(2n, 2n + 1)) for some n.

Now if the matrices are finite, then there is bound to be a cycle. Perhaps such a cycle will reach a Nash equilibrium and then stabilize. Perhaps it will just go on and on, with immediately succeeding levels of knowledge giving different payoffs, with the player who knows a bit more having an advantage.

Further Work and Open Questions

In this paper we have looked only at levels of knowledge for single formulas. However, levels of knowledge for related formulas may be connected. For example, if $A = B \lor C$, then $L(B, s) \cup L(C, s) \subseteq L(A, s)$. So one could ask, given the Lindenbaum algebra \mathcal{A} of ground formulas and the Boolean algebra \mathcal{B} of subsets of Σ_c^* , which maps from \mathcal{A} to \mathcal{B} can arise as level maps? We know that the maps must preserve order and that the images must be regular, downward closed sets, but what more can we show?

A second direction of inquiry is to ask how actual game playing and knowledge interact. We have shown what sorts of levels can arise and shown that they are relevant to group strategies as well as to individual strategies within groups. But clearly much more needs to be done.

A final line of research is to bring the current work into closer contact with a lot of other work on knowledge revision which begins with Plaza and proceeds through [Ger], [BMP], [Dit]. We also need to relate the work with the work of Stalnaker on models of knowledge where probabilities are taken into account. **Acknowledgements:** We thank Steven Brams, Eva Cogan, Joe Halpern, Karen Kletter, Eric Pacuit, Debraj Ray and two referees for very useful comments.

References:

pp. 269-280.

[AS] R. Aumann and S. Sorin, Cooperation and bounded recall, Games and Economic Behavior 1 (1989) 5-39.

[Ba] J. Barwise, Three views of common knowledge, in *TARK-2*, Ed. M. Vardi, Morgan Kaufmann 1988, pp. 369-380.

[BMP] A. Baltag, L. Moss and S. Solecki, The logic of public announcements, common knowledge and private suspicions, Technical report #238, Indiana University Cognitive Science Program (earlier version appeared in *TARK 1998*)

[Chw] M. Chwe, Rational Ritual, Princeton U. Press, 2001

[CM2] H. H. Clark and C. R. Marshall, Definite reference and mutual knowledge, in *Elements of Discourse Understanding*, Ed. Joshi, Webber and Sag, Cambridge U. Press, 1981.

[Dit] H. van Ditmarsch, *Knowledge Games*, Doctoral dissertation, U. of Groeningen (2000).

[FHMV] R. Fagin, J. Halpern, Y.Moses and M. Vardi, *Reasoning abour Knowledge*, MIT press, 1995.

[Ger] J. Gerbrandy, Dynamic epistemic logic, in *Logic, Language and Information II*, CSLI press 1999.

[Hi] G. Higman, Ordering by divisibility in abstract algebras, Proc. London Math. Soc.3 (1952) 326-336

[HM] J. Halpern and Y. Moses, Knowledge and Common Knowledge in a Distributed Environment, *Proc. 3rd ACM Symposium on Distributed Computing* 1984 pp. 50-61
[HZ] J. Halpern and L. Zuck, A little knowledge goes a long way, *Proc. 6th PODC*, 1987,

32

[JP] D.H.J. de Jongh and R. Parikh, Well partial orderings and hierarchies, Proc. Kon. Ned. Akad. Sci Series A 80 (1977) 195- 207.

[Lew] D. Lewis, Convention, a Philosophical Study, Harvard U. Press, 1969.

[MGM] R. Marvin, M. Greenberg and D. Mossler, The Early development of conceptual perspective thinking, *Child Development*, **47** (1976) 511-514.

[MT] Y. Moses and M. Tuttle, Programming simultaneous actions using common knowledge, Research Report MIT/LCS/TR-369 (1987)

[Pa1] R. Parikh, Knowledge and the problem of logical omniscience, *ISMIS-87*, North Holland, pp. 432-439.

[Pa2] R. Parikh, Finite and infinite dialogues, Proceedings of a Workshop on Logic and Computer Science, ed. Moschovakis, Sprinter 1991, 481-98.

[Pa3] R. Parikh, Social software, *Synthese*, **132**, Sep 2002, 187-211.

[PK] R. Parikh and P. Krasucki, Levels of knowledge in distributed computing, Sadhana
– Proc. Ind. Acad. Sci. 17 (1992) pp. 167-191.

[Pl] J. Plaza, Logics of public announcements, Proceedings 4th International Symposium on Methodologies for Intelligent Systems, 1989.

[PR] R. Parikh and R. Ramanujam, Distributed computing and the logic of knowledge, Logics of Programs 1985, Springer LNCS 193, 256-268.

[SA] R. Schank and R. Abelson, *Scripts, Plans, Goals, and Understanding*, Erlbaum Hillsdale, NJ (1977)

[SP] C. Steinsvold and R. Parikh, A Modal analysis of some phenomena in child psychology, *Bulletin of Symbolic Logic*, Mar 2002, Logic Colloquium '01, page 158.

[Sta] R. Stalnaker, Knowledge, belief and counterfactual reasoning in games, in *The Logic of Strategy*, Ed. Bicchieri et al, Oxford University Press, 1999.

[St] C. Steinsvold, Trust and other modal phenomena, research report, CUNY Graduate Center, February 2002.

[WP] H. Wimmer and J. Perner, Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception, *Cognition*, **13** (1983) 103-128.