Modal Logic
Common Knowledge and Agreement
Theorems

Eric Pacuit
University of Maryland, College Park
ai.stanford.edu/~epacuit

April 19, 2012
Plan

- Common Knowledge
- Agreeing to Disagree
“Common Knowledge” is informally described as what any fool would know, given a certain situation: It encompasses what is relevant, agreed upon, established by precedent, assumed, being attended to, salient, or in the conversational record.
“Common Knowledge” is informally described as what any fool would know, given a certain situation: It encompasses what is relevant, agreed upon, established by precedent, assumed, being attended to, salient, or in the conversational record.

It is not Common Knowledge who “defined” Common Knowledge!
The first formal definition of common knowledge?


The first formal definition of common knowledge?

The first rigorous analysis of common knowledge
The first formal definition of common knowledge?

The first rigorous analysis of common knowledge

**Fixed-point definition**: $\gamma := i \text{ and } j \text{ know that } (\varphi \text{ and } \gamma)$

The first formal definition of common knowledge?

The first rigorous analysis of common knowledge

**Fixed-point definition:** \( \gamma := i \text{ and } j \text{ know that (} \varphi \text{ and } \gamma \text{)} \)

**Shared situation:** There is a *shared situation* \( s \) such that (1) \( s \) entails \( \varphi \), (2) \( s \) entails everyone knows \( \varphi \), plus other conditions
The “Standard” Account


The “Standard” Account

$W$ is a set of states or worlds.
The “Standard” Account

An event/proposition is any (definable) subset $E \subseteq W$
The “Standard” Account

At each state, agents are assigned a set of states they consider possible (according to their information). The information may be (in)correct, partitional, ....
The “Standard” Account

Knowledge Function: $K_i : \wp(W) \rightarrow \wp(W)$ where $K_i(E) = \{ w \mid R_i(w) \subseteq E \}$
The “Standard” Account

\[ w \in K_A(E) \text{ and } w \notin K_B(E) \]
The “Standard” Account

The model also describes the agents’ higher-order knowledge/beliefs
The “Standard” Account

Everyone Knows: \[ K(E) = \bigcap_{i \in A} K_i(E), \quad K^0(E) = E, \quad K^m(E) = K(K^{m-1}(E)) \]
The “Standard” Account

**Common Knowledge:** \( C : \wp(W) \to \wp(W) \) with

\[
C(E) = \bigcap_{m \geq 0} K^m(E)
\]
The “Standard” Account

\[ w \in K(E) \quad w \notin C(E) \]
The “Standard” Account

\[ w \in C(E) \]
**Fact.** For all $i \in A$ and $E \subseteq W$, $K_i C(E) = C(E)$. 
Fact. For all $i \in \mathcal{A}$ and $E \subseteq W$, $K_i C(E) = C(E)$.

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)
Fact. For all $i \in A$ and $E \subseteq W$, $K_i C(E) = C(E)$.

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

**Fact.** An event $E$ is commonly known iff some self-evident event that entails $E$ obtains.
**Fact.** For all $i \in \mathcal{A}$ and $E \subseteq W$, $K_i C(E) = C(E)$.

An event $F$ is **self-evident** if $K_i(F) = F$ for all $i \in \mathcal{A}$.

**Fact.** An event $E$ is commonly known iff some self-evident event that entails $E$ obtains.

**Fact.** $w \in C(E)$ if every finite path starting at $w$ ends in a state in $E$.

The following axiomatize common knowledge:

- $C(\varphi \rightarrow \psi) \rightarrow (C\varphi \rightarrow C\psi)$
- $C\varphi \rightarrow (\varphi \land EC\varphi)$ (Fixed-Point)
- $C(\varphi \rightarrow E\varphi) \rightarrow (\varphi \rightarrow C\varphi)$ (Induction)
An Example

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.
An Example

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).
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?
An Example

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?
Agreeing to Disagree

**Theorem:** Suppose that $n$ agents share a common prior and have different private information. If there is common knowledge in the group of the posterior probabilities, then the posteriors must be equal.

Agreeing to Disagree

**Theorem:** Suppose that $n$ agents share a common prior and have different private information. If there is common knowledge in the group of the posterior probabilities, then the posteriors must be equal.

2 Scientists Perform an Experiment

They agree the true state is one of seven different states.
2 Scientists Perform an Experiment

They agree on a common prior.
2 Scientists Perform an Experiment

They agree that Experiment 1 would produce the blue partition.
They agree that Experiment 1 would produce the blue partition and Experiment 2 the red partition.
They are interested in the truth of $E = \{w_2, w_3, w_5, w_6\}$. 
So, they agree that $P(E) = \frac{24}{32}$. 
Also, that if the true state is \( w_1 \), then Experiment 1 will yield

\[
P(E|I) = \frac{P(E \cap I)}{P(I)} = \frac{12}{14}
\]
Suppose the true state is $w_7$ and the agents preform the experiments.
Suppose the true state is $w_7$, then $Pr_1(E) = \frac{12}{14}$
2 Scientists Perform an Experiment

Then $Pr_1(E) = \frac{12}{14}$ and $Pr_2(E) = \frac{15}{21}$
Suppose they exchange emails with the new subjective probabilities: $Pr_1(E) = \frac{12}{14}$ and $Pr_2(E) = \frac{15}{21}$.
2 Scientists Perform an Experiment

Agent 2 learns that $w_4$ is NOT the true state (same for Agent 1).
Agent 2 learns that $w_4$ is **NOT** the true state (same for Agent 1).
Agent 1 learns that $w_5$ is **NOT** the true state (same for Agent 1).
The new probabilities are $Pr_1(E|I') = \frac{7}{9}$ and $Pr_2(E|I') = \frac{15}{17}$
After exchanging this information ($Pr_1(E|I') = \frac{7}{9}$ and $Pr_2(E|I') = \frac{15}{17}$), Agent 2 learns that $w_3$ is NOT the true state.
2 Scientists Perform an Experiment

No more revisions are possible and the agents agree on the posterior probabilities.
Dissecting Aumann’s Theorem

- Qualitative versions: like-minded individuals cannot agree to make different decisions.


The Framework

Knowledge Structure: \( \langle W, \{\Pi_i\}_{i \in A} \rangle \) where each \( \Pi_i \) is a partition on \( W \) (\( \Pi_i(w) \) is the cell in \( \Pi_i \) containing \( w \)).

Decision Function: Let \( D \) be a nonempty set of decisions. A decision function for \( i \in A \) is a function \( d_i : W \rightarrow D \). A vector \( d = (d_1, \ldots, d_n) \) is a decision function profile. Let \( [d_i = d] = \{w \mid d_i(w) = d\} \).
The Framework

**Knowledge Structure:** $\langle W, \{\Pi_i\}_{i \in A}\rangle$ where each $\Pi_i$ is a partition on $W$ ($\Pi_i(w)$ is the cell in $\Pi_i$ containing $w$).

**Decision Function:** Let $D$ be a nonempty set of decisions. A decision function for $i \in A$ is a function $d_i : W \rightarrow D$. A vector $d = (d_1, \ldots, d_n)$ is a decision function profile. Let $[d_i = d] = \{w \mid d_i(w) = d\}$.

(A1) Each agent knows her own decision:

$$[d_i = d] \subseteq K_i([d_i = d])$$
Comparing Knowledge

\([j \succeq i]\): agent \(j\) is at least as knowledgeable as agent \(i\).

\[
[j \succeq i] := \bigcap_{E \in \wp(W)} (K_i(E) \Rightarrow K_j(E)) = \bigcap_{E \in \wp(W)} (\neg K_i(E) \cup K_j(E))
\]
Comparing Knowledge

\[ j \succeq i \]: agent \( j \) is at least as knowledgeable as agent \( i \).

\[
[j \succeq i] := \bigcap_{E \in \mathcal{P}(W)} (K_i(E) \Rightarrow K_j(E)) = \bigcap_{E \in \mathcal{P}(W)} (\neg K_i(E) \cup K_j(E))
\]

\( w \in [j \succeq i] \) then \( j \) knows at \( w \) every event that \( i \) knows there.
Comparing Knowledge

\[ [j \succeq i] : \text{agent } j \text{ is at least as knowledgeable as agent } i. \]

\[ [j \succeq i] := \bigcap_{E \in \wp(W)} (K_i(E) \Rightarrow K_j(E)) = \bigcap_{E \in \wp(W)} (\neg K_i(E) \cup K_j(E)) \]

\[ w \in [j \succeq i] \text{ then } j \text{ knows at } w \text{ every event that } i \text{ knows there.} \]

\[ [j \sim i] = [j \succeq i] \cap [i \succeq j] \]
Interpersonal Sure-Thing Principle (ISTP)

For any pair of agents $i$ and $j$ and decision $d$,

$$K_i([j \succeq i] \cap [d_j = d]) \subseteq [d_i = d]$$
Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case.
Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case. If they are exposed to different evidence, they may reach different decisions.
Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case. If they are exposed to different evidence, they may reach different decisions. Yet, being the students of the same academy, the method by which they arrive at their conclusions is the same.
Interpersonal Sure-Thing Principle (ISTP): Illustration

Suppose that Alice and Bob, two detectives who graduated the same police academy, are assigned to investigate a murder case. If they are exposed to different evidence, they may reach different decisions. Yet, being the students of the same academy, the method by which they arrive at their conclusions is the same. Suppose now that detective Bob, a father of four who returns home every day at five o'clock, collects all the information about the case at hand together with detective Alice.
Interpersonal Sure-Thing Principle (ISTP): Illustration

However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob.
However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob. Obviously, Bob knows that Alice is at least as knowledgeable as he is.
However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob. Obviously, Bob knows that Alice is at least as knowledgeable as he is. Suppose that he also knows what Alice's decision is.
Interpersonal Sure-Thing Principle (ISTP): Illustration

However, Alice, single and a workaholic, continues to collect more information every day until the wee hours of the morning — information which she does not necessarily share with Bob. Obviously, Bob knows that Alice is at least as knowledgeable as he is. Suppose that he also knows what Alice’s decision is. Since Alice uses the same investigation method as Bob, he knows that had he been in possession of the more extensive knowledge that Alice has collected, he would have made the same decision as she did. Thus, this is indeed his decision.
Implications of ISTP

**Proposition.** If the decision function profile $d$ satisfies ISTP, then

$$[i \sim j] \subseteq \bigcup_{d \in D} ([d_i = d] \cap [d_j = d])$$
Agent $i$ is an **epistemic dummy** if it is always the case that all the agents are at least as knowledgeable as $i$. That is, for each agent $j$,

$$[j \succeq i] = W$$

A decision function profile $d$ on $\langle W, \Pi_1, \ldots, \Pi_n \rangle$ is **ISTP expandable** if for any expanded structure $\langle W, \Pi_1, \ldots, \Pi_{n+1} \rangle$ where $n + 1$ is an epistemic dummy, there exists a decision function $d_{n+1}$ such that $(d_1, d_2, \ldots, d_{n+1})$ satisfies ISTP.
Suppose that after making their decisions, Alice and Bob are told that another detective, one E.P. Dummy, who graduated the very same police academy, had also been assigned to investigate the same case.
Suppose that after making their decisions, Alice and Bob are told that another detective, one E.P. Dummy, who graduated the very same police academy, had also been assigned to investigate the same case. In principle, they would need to review their decisions in light of the third detectives knowledge: knowing what they know about the third detective, his usual sources of information, for example, may impinge upon their decision.
But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set.
But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set. Thus, it is commonly known that every detective is at least as knowledgeable as Dummy.
ISTP Expandability: Illustration

But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set. Thus, it is commonly known that every detective is at least as knowledgeable as Dummy. The news that he had been assigned to the same case is completely irrelevant to the conclusions that Alice and Bob have reached. Obviously, based on the information he gets from the media, Dummy also makes a decision.
But this is not so in the case of detective Dummy. It is commonly known that the only information source of this detective, known among his colleagues as the couch detective, is the TV set. Thus, it is commonly known that every detective is at least as knowledgeable as Dummy. The news that he had been assigned to the same case is completely irrelevant to the conclusions that Alice and Bob have reached. Obviously, based on the information he gets from the media, Dummy also makes a decision. We may assume that the decisions made by the three detectives satisfy the ISTP, for exactly the same reason we assumed it for the two detectives decisions.
Generalized Agreement Theorem

If \( d \) is an ISTP expandable decision function profile on a partition structure \( \langle W, \Pi_1, \ldots, \Pi_n \rangle \), then for any decisions \( d_1, \ldots, d_n \) which are not identical, \( C(\bigcap_i [d_i = d_i]) = \emptyset \).
Common $p$-belief

**Theorem.** If the posteriors of an event $X$ are common $p$-belief at some state $w$, then any two posteriors can differ by at most $2(1 - p)$.

Analyzing Agreement Theorems in Dynamic Epistemic/Doxastic Logic


Next: Dynamics of Knowledge and Belief