Rationality

Lecture 11

Eric Pacuit

Center for Logic and Philosophy of Science
Tilburg University
ai.stanford.edu/∼epacuit
e.j.pacuit@uvt.nl

May 2, 2011

Equilibrium Selection Problem

		Bob D	Bob C
Ann	D	1,1	0,0
	C	0,0	1,1

What should/will Ann (Bob) do?

Equilibrium Selection Problem

		Bob	
		D	C
Ann	D	1,1	0,0
	C	0,0	1,1

What should/will Ann (Bob) do?

Equilibrium Selection Problem

Ann \ Bob	D	C
D	3,3	0,0
C	0,0	1,1

What should/will Ann (Bob) do?

Footballer Example

A and B are players in the same football team. A has the ball, but an opposing player is converging on him. He can pass the ball to B, who has a chance to shoot. There are two directions in which A can move the ball, *left* and *right*, and correspondingly, two directions in which B can run to intercept the pass. If both choose *left* there is a 10% chance that a goal will be scored. If they both choose *right*, there is a 11% change. Otherwise, the chance is zero. There is no time for communication; the two players must act simultaneously.

What should they do?

R. Sugden. *The Logic of Team Reasoning*. Philosophical Explorations (6)3, pgs. 165 - 181 (2003).

Footballer Example

		B l	B r
A	l		
A	r		

Footballer Example

		B l	B r
A	l	10,10	0,0
A	r	0,0	11,11

A: What should I do?

Footballer Example

		B l	B r
A	l	10,10	0,0
A	r	0,0	11,11

A: What should I do? r if the probability of B choosing r is $> \frac{10}{21}$ and l if the probability of B choosing l is $> \frac{11}{21}$
(symmetric reasoning for B)

Footballer Example

			B	
			l	r
A		l	10,10	0,0
		r	0,0	11,11

A: What should I do? r if the probability of B choosing r is $> \frac{10}{21}$ and l if the probability of B choosing l is $> \frac{11}{21}$
(symmetric reasoning for B)

Footballer Example

		B	
		l	r
A	l	10,10	0,0
	r	0,0	11,11

Rationality in Interaction

What does it mean to be rational when the outcome of one's action depends upon the actions of other people and everyone is trying to guess what the others will do?

Rationality in Interaction

What does it mean to be rational when the outcome of one's action depends upon the actions of other people and everyone is trying to guess what the others will do?

In social interaction, rationality has to be enriched with further assumptions about individuals' **mutual knowledge and beliefs**, *but these assumptions are not without consequence.*

C. Bicchieri. *Rationality and Game Theory*. Chapter 10 in [HR].

Example: Common Knowledge

Suppose there are two friends Ann and Bob are on a bus separated by a crowd.

Example: Common 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?".

Example: Common 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.

Example: Common 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?

Example: Common 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.

"*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?

M. Friedell. *On the Structure of Shared Awareness*. Behavioral Science (1969).

R. Aumann. *Agreeing to Disagree*. Annals of Statistics (1976).

The first formal definition of common knowledge?

M. Friedell. *On the Structure of Shared Awareness*. Behavioral Science (1969).

R. Aumann. *Agreeing to Disagree*. Annals of Statistics (1976).

The first rigorous analysis of common knowledge

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

The first formal definition of common knowledge?

M. Friedell. *On the Structure of Shared Awareness*. Behavioral Science (1969).

R. Aumann. *Agreeing to Disagree*. Annals of Statistics (1976).

The first rigorous analysis of common knowledge

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

Fixed-point definition: $\gamma :=$ *i* and *j* know that (φ and γ)

G. Harman. *Review of* Linguistic Behavior. Language (1977).

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

The first formal definition of common knowledge?

M. Friedell. *On the Structure of Shared Awareness*. Behavioral Science (1969).

R. Aumann. *Agreeing to Disagree*. Annals of Statistics (1976).

The first rigorous analysis of common knowledge

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

Fixed-point definition: $\gamma :=$ i and j know that (φ and γ)

G. Harman. *Review of* Linguistic Behavior. Language (1977).

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

Shared situation: There is a *shared situation* s such that (1) s entails φ, (2) s entails everyone knows φ, plus other conditions

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

M. Gilbert. *On Social Facts*. Princeton University Press (1989).

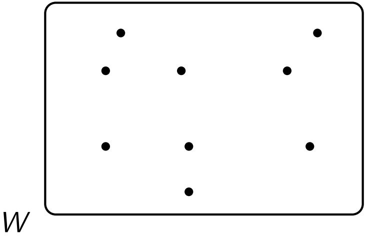
P. Vanderschraaf and G. Sillari. *"Common Knowledge", The Stanford Encyclopedia of Philosophy (2009)*. http://plato.stanford.edu/entries/common-knowledge/.

The "Standard" Account

R. Aumann. *Agreeing to Disagree*. Annals of Statistics (1976).

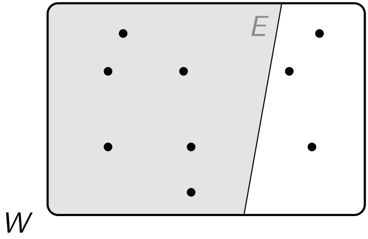
R. Fagin, J. Halpern, Y. Moses and M. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

The “Standard” Account



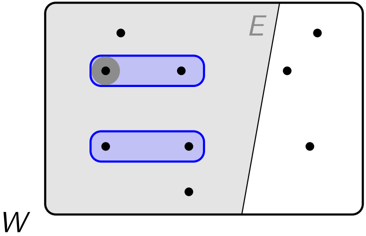
W is a set of **states** or **worlds**.

The "Standard" Account



An **event**/**proposition** is any (definable) subset $E \subseteq W$

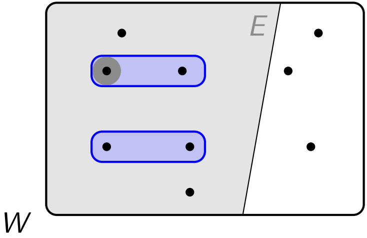
The “Standard” Account



At each state, agents are assigned a set of states they *consider possible* (according to their information).

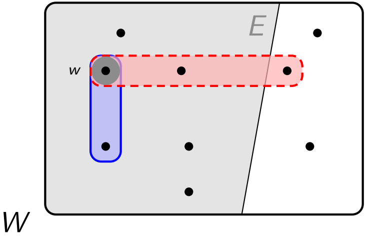
The information may be (in)correct, partitional, ...

The "Standard" Account



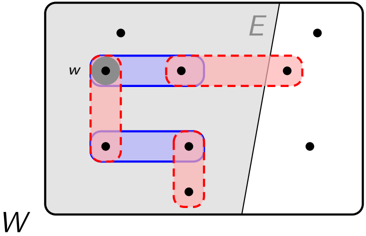
Knowledge Function: $K_i : \wp(W) \rightarrow \wp(W)$ where $K_i(E) = \{w \mid R_i(w) \subseteq E\}$

The "Standard" Account



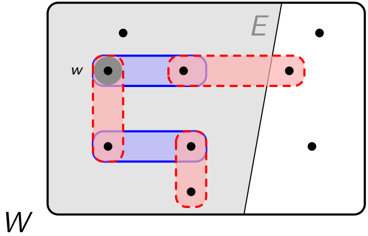
$w \in K_A(E)$ and $w \notin K_B(E)$

The “Standard” Account



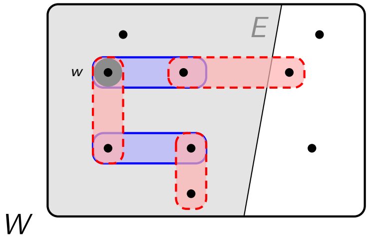
The model also describes the agents’ **higher-order knowledge/beliefs**

The “Standard” Account



Everyone Knows: $K(E) = \bigcap_{i\in\mathcal{A}} K_i(E)$, $K^0(E) = E$, $K^m(E) = K(K^{m-1}(E))$

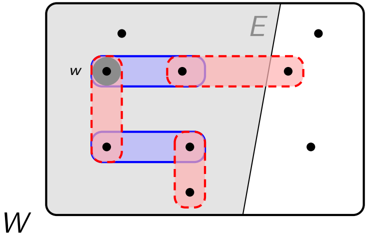
The "Standard" Account



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

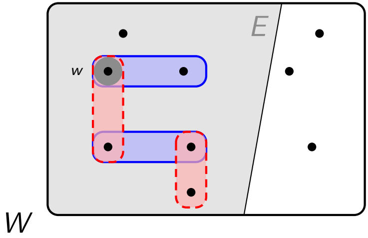
$$C(E) = \bigcap_{m \geq 0} K^m(E)$$

The "Standard" Account



$w \in K(E) \qquad w \notin C(E)$

The "Standard" Account



$w \in C(E)$

Fact. For all $i \in \mathcal{A}$ 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 $i \in \mathcal{A}$ 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. 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:

- $C(\varphi \to \psi) \to (C\varphi \to C\psi)$
- $C\varphi \to (\varphi \wedge EC\varphi)$ (Fixed-Point)
- $C(\varphi \to E\varphi) \to (\varphi \to C\varphi)$ (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).

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?

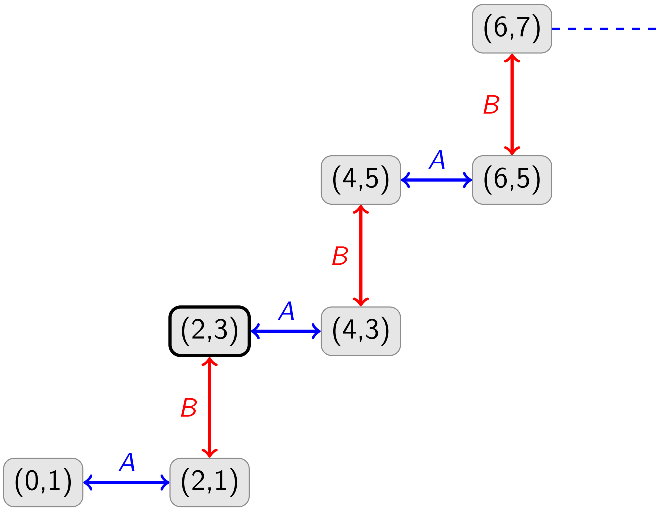
An Example

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?


(6,7)
B
(4,5)
A
(6,5)
B
(2,3)
A
(4,3)
B
(0,1)
A
(2,1)

Some Issues

Some Issues

- What *does* a group know/believe/accept? vs. what *can* a group (come to) know/believe/accept?

Some Issues

- What *does* a group know/believe/accept? vs. what *can* a group (come to) know/believe/accept?

C. List. *Group knowledge and group rationality: a judgment aggregation perspective*. Episteme (2008).

Some Issues

- What *does* a group know/believe/accept? vs. what *can* a group (come to) know/believe/accept?

C. List. *Group knowledge and group rationality: a judgment aggregation perspective*. Episteme (2008).

- Other "group informational attitudes": distributed knowledge, common belief, . . .

Some Issues

- What *does* a group know/believe/accept? vs. what *can* a group (come to) know/believe/accept?

C. List. *Group knowledge and group rationality: a judgment aggregation perspective*. Episteme (2008).

- Other "group informational attitudes": distributed knowledge, common belief, ...

- Common knowledge/belief of *rationality*

Some Issues

- What *does* a group know/believe/accept? vs. what *can* a group (come to) know/believe/accept?

C. List. *Group knowledge and group rationality: a judgment aggregation perspective*. Episteme (2008).

- Other "group informational attitudes": distributed knowledge, common belief, ...

- Common knowledge/belief of *rationality*

- Where does common knowledge come from?

Key Assumptions

CK1 The structure of the game, including players' strategy sets and payoff functions, is common knowledge among the players.

CK2 The players are rational (i.e., they are expected utility maximizers) and this is common knowledge.

Common Knowledge of Rationality: Iterated Removal of Strictly Dominated Strategies

Bob

Ann		L	R
	U	1,2	0,1
	D	0,1	1,0

Common Knowledge of Rationality: Iterated Removal of Strictly Dominated Strategies

		Bob	
		L	R
Ann	U	1,2	0,1
	D	0,1	1,0

There is no prior such that R is rational for Bob.

Common Knowledge of Rationality: Iterated Removal of Strictly Dominated Strategies

		Bob	
		L	R
Ann	U	1,2	0,1
	D	0,1	1,0

If Ann knows this, then she does not consider R a option for Bob

Common Knowledge of Rationality: Iterated Removal of Strictly Dominated Strategies

Ann / Bob

	L	R
U	1,2	0,1
D	0,1	1,0

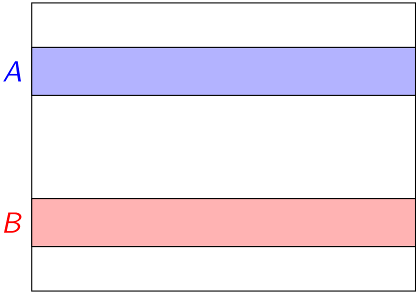
So, U is the only rational choice.

Common knowledge of rationality (players will not choose strictly dominated actions) leads to a process of iterated removal of strictly dominated strategies.

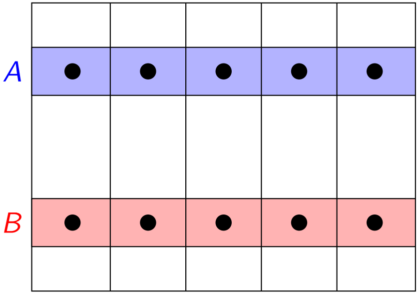
Common knowledge of rationality (players will not choose strictly dominated actions) leads to a process of iterated removal of strictly dominated strategies.

What about *weak dominance*?

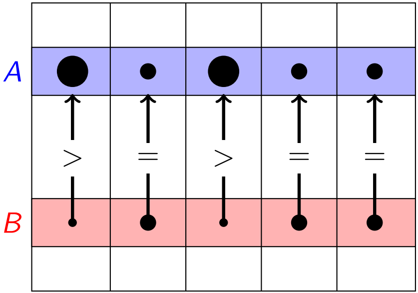
Weak Dominance



Weak Dominance



Weak Dominance



Iterated Admissibility

	L	R
U	1,1	0,1
D	0,2	1,0

Iterated Admissibility

	L	R
U	1,1	0,1
D	0,2	1,0

Suppose rationality incorporates *weak dominance* (i.e., *admissibility* or *cautiousness*).

Iterated Admissibility

	L	R
U	1,1	0,1
D	0,2	1,0

Suppose rationality incorporates *weak dominance* (i.e., *admissibility* or *cautiousness*).

1. Both Row and Column should use a *full-support* probability measure
2. But if Row thinks that Column is **rational** then should she not assign probability 1 to L?

Iterated Admissibility

	L	R
U	1,1	0,1
D	0,2	1,0

Suppose rationality incorporates *weak dominance* (i.e., *admissibility* or *cautiousness*).

1. Both Row and Column should use a *full-support* probability measure
2. But if Row thinks that Column is **rational** then should she not assign probability 1 to L?

The condition that the players incorporate admissibility *into their rationality calculations seems to conflict with the condition that the players think the other players are rational (there is a tension between admissibility and strategic reasoning)*

Iterated Removal of Weakly Dominated Strategies

Bob

Ann		L	R
	T	1,1	1,0
	B	1,0	0,1

Iterated Removal of Weakly Dominated Strategies

Ann / Bob

	L	R
T	1,1	1,0
B	1,0	0,1

T weakly dominates B

Iterated Removal of Weakly Dominated Strategies

		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1

Then L strictly dominates R.

Iterated Removal of Weakly Dominated Strategies

Bob

Ann		L	R
	T	1,1	1,0
	B	1,0	0,1

The IA set

Iterated Removal of Weakly Dominated Strategies

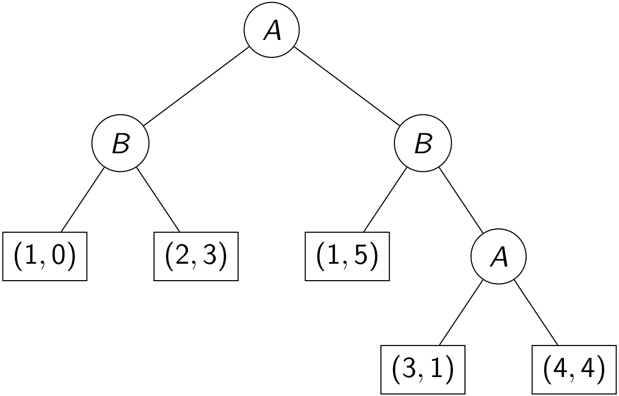
Bob

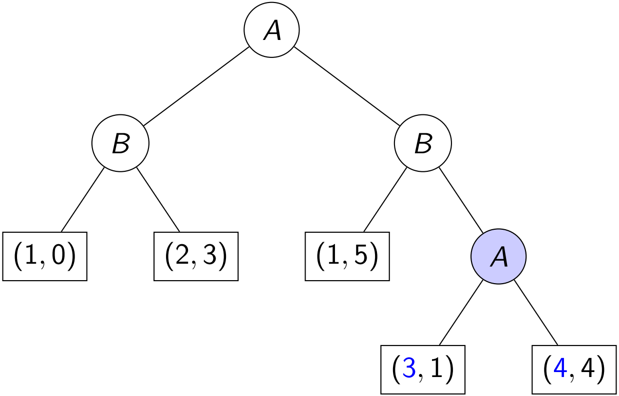
Ann		L	R
	T	1,1	1,0
	B	1,0	0,1

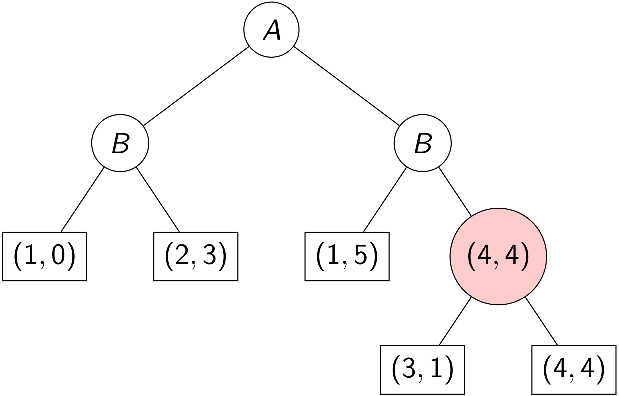
But, now what is the reason for not playing B?

Backwards Induction

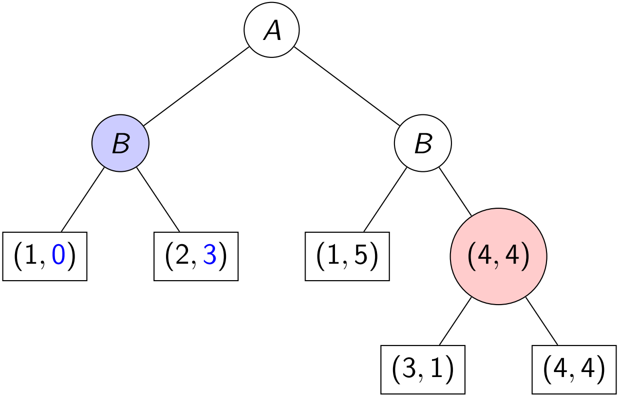
Invented by Zermelo, Backwards Induction is an iterative algorithm for "solving" and extensive game.

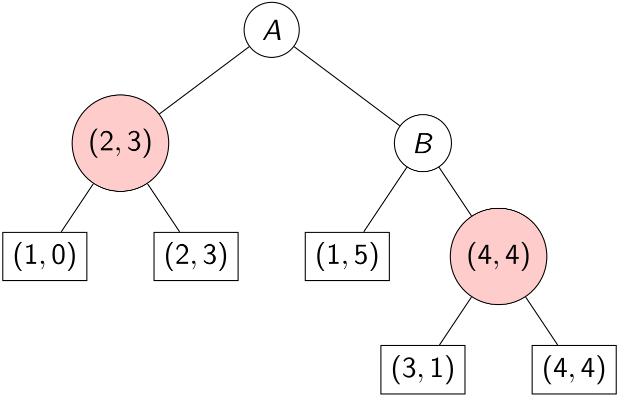

A
B
B
(1, 0)
(2, 3)
(1, 5)
A
(3, 1)
(4, 4)


A
B
B
(1, 0)
(2, 3)
(1, 5)
A
(3, 1)
(4, 4)

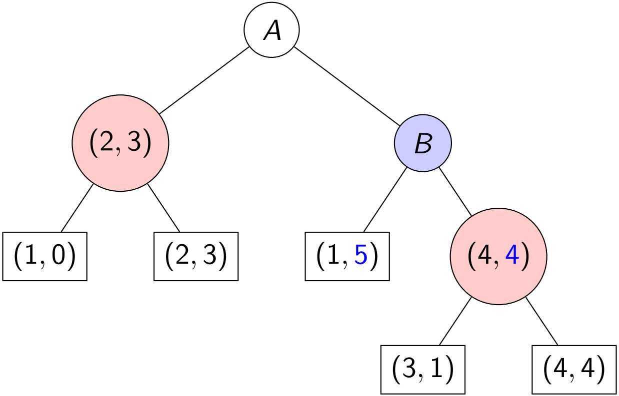


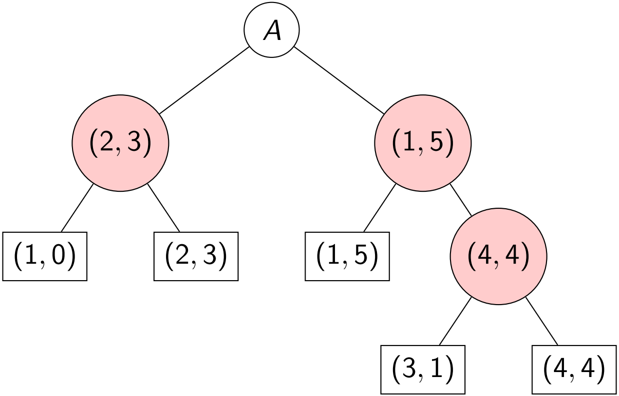
A
B
B
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

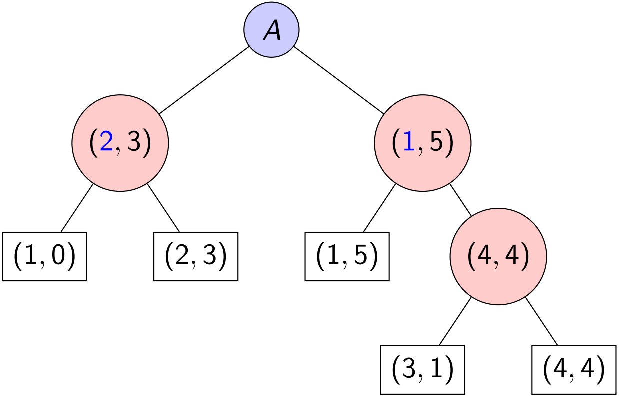

A
B
B
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

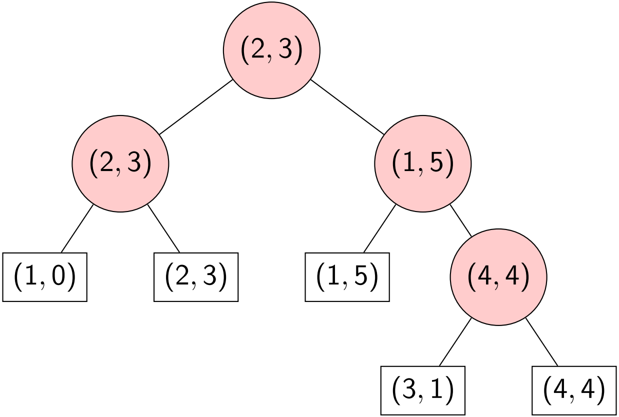


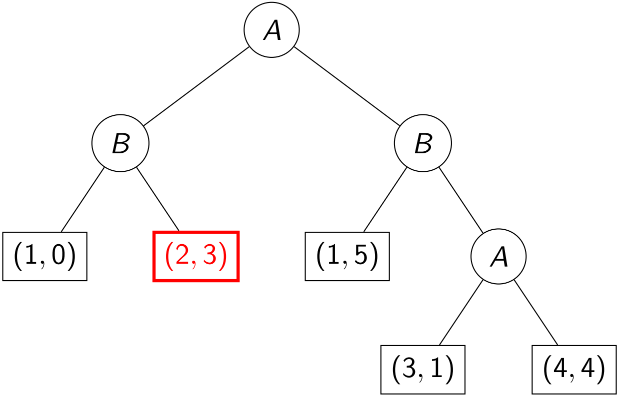
A
$(2,3)$
B
$(1,0)$
$(2,3)$
$(1,5)$
$(4,4)$
$(3,1)$
$(4,4)$

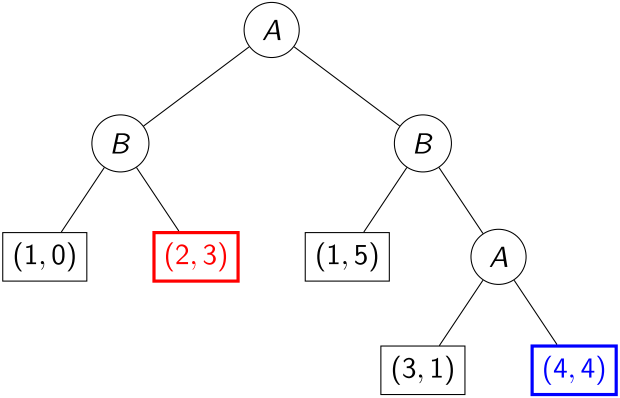

A
(2, 3)
B
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)


A
(2, 3)
(1, 5)
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

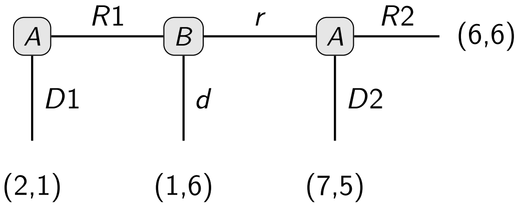

A
(2, 3)
(1, 5)
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)


(2, 3)
(2, 3)
(1, 5)
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

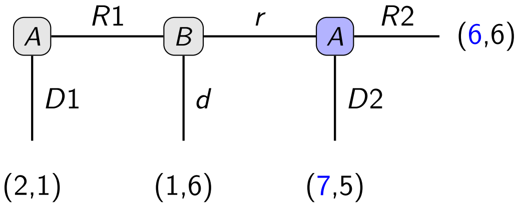

A
B
B
(1, 0)
(2, 3)
(1, 5)
A
(3, 1)
(4, 4)


A
B
B
(1, 0)
(2, 3)
(1, 5)
A
(3, 1)
(4, 4)

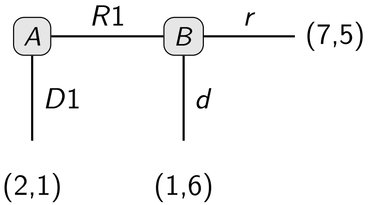
BI Puzzle



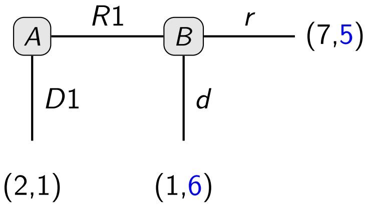
BI Puzzle



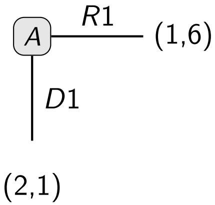
BI Puzzle



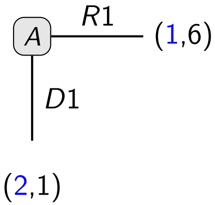
BI Puzzle



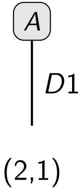
BI Puzzle



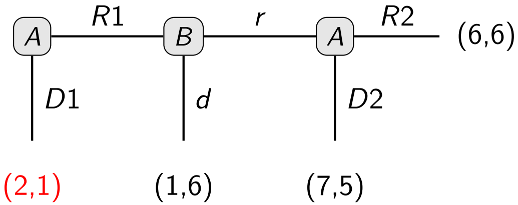
BI Puzzle



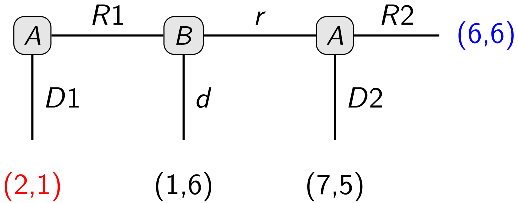
BI Puzzle



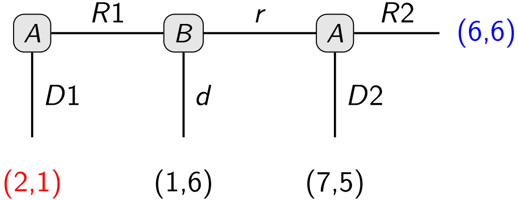
BI Puzzle



But what if...



But what if...



- Are the players *irrational*?
- What argument leads to the BI solution?

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

$\cdots$

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

What about "tit-for-tat"?

Repeated Prisoner's Dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

	C	D
C	3,3	0,4
D	4,0	1,1

What about "tit-for-tat"?

Is anything missing in these models?

Formally, a game is described by its strategy sets and payoff functions.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12% respectively.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12% respectively. But the political situations are quite different.

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12% respectively. But the political situations are quite different. The difference lies in the attitudes of the players, in their expectations about each other, in custom, and in history, though the rules of the game do not distinguish between the two situations.

R. Aumann and J. H. Dreze. *Rational Expectation in Games*. American Economic Review (2008).

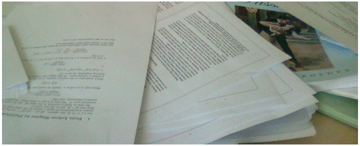
Two questions

- What should the players *do* in a game-theoretic situation and what should they expect? (Assuming everyone is **rational** and recognize each other's rationality)

- What are the assumptions about rationality and the players' knowledge/beliefs underlying the various solution concepts? *Why* would the agents' follow a particular solution concept?

Writing a paper together

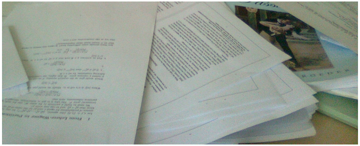
	C	D
C		
D		



Writing a paper together

Problem of Cooperation.

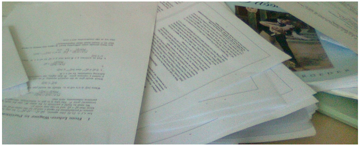
	C	D
C	3,3	0,4
D	4,0	1,1



Writing a paper together

Problem of Coordination.

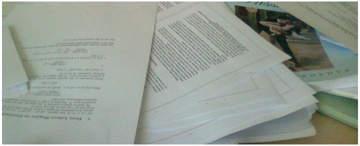
	C	D
C	3,3	0,0
D	0,0	1,1



Writing a paper together

	C	D
C		
D		

Intuitively, we solve these problem by working together.
This is the question of collective agency.



R. Cubitt and R. Sugden. *Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory*. Economics and Philosophy, 19, pgs. 175-210 , 2003..

Reason to Believe

$B_i\varphi$: "i believes φ"

Reason to Believe

$B_i\varphi$: "i believes φ" vs. $R_i(\varphi)$: "i has a *reason to believe* φ"

Reason to Believe

$B_i\varphi$: "i believes φ" vs. $R_i(\varphi)$: "i has a *reason to believe* φ"

- "Although it is an essential part of Lewis' theory that human beings are *to some degree* rational, he does not want to make the strong rationality assumptions of conventional decision theory or game theory." (CS, pg. 184).

Reason to Believe

$B_i\varphi$: "i believes φ" vs. $R_i(\varphi)$: "i has a *reason to believe* φ"

- "Although it is an essential part of Lewis' theory that human beings are *to some degree* rational, he does not want to make the strong rationality assumptions of conventional decision theory or game theory." (CS, pg. 184).
- Anyone who accept the rules of arithmetic has a reason to believe $618 \times 377 = 232,986$, but most of us do not hold have firm beliefs about this.

Reason to Believe

$B_i\varphi$: "i believes φ" vs. $R_i(\varphi)$: "i has a *reason to believe* φ"

- "Although it is an essential part of Lewis' theory that human beings are *to some degree* rational, he does not want to make the strong rationality assumptions of conventional decision theory or game theory." (CS, pg. 184).
- Anyone who accept the rules of arithmetic has a reason to believe $618 \times 377 = 232,986$, but most of us do not hold have firm beliefs about this.
- Definition: $R_i(\varphi)$ means φ is true within some logic of reasoning that is *endorsed* by (that is, accepted as a normative standard by) person i...φ must be either regarded as *self-evident* or derivable by rules of inference (deductive or inductive)

A indicates to i that φ

A is a "state of affairs"

$A\ ind_i\ \varphi$: i's reason to believe that A holds *provides* i's reason for believing that φ is true.

(A1)For all i, for all A, for all φ: $[R_i(A\ holds) \wedge (A\ ind_i\ \varphi)] \Rightarrow R_i(\varphi)$

Some Properties

Some Properties

- $[(A \ holds) \ entails \ (A' \ holds)] \Rightarrow A \ ind_i(A' \ holds)$

Some Properties

- $[(A \textit{ holds}) \textit{ entails } (A' \textit{ holds})] \Rightarrow A \textit{ ind}_i(A' \textit{ holds})$
- $[(A \textit{ ind}_i\ \varphi) \wedge (A \textit{ ind}_i \psi)] \Rightarrow A \textit{ ind}_i(\varphi \wedge \psi)$

Some Properties

- $[(A\ \mathit{holds})\ \mathit{entails}\ (A'\ \mathit{holds})] \Rightarrow A\ \mathit{ind}_i(A'\ \mathit{holds})$
- $[(A\ \mathit{ind}_i\ \varphi) \wedge (A\ \mathit{ind}_i \psi)] \Rightarrow A\ \mathit{ind}_i(\varphi \wedge \psi)$
- $[(A\ \mathit{ind}_i[A'\ \mathit{holds}]) \wedge (A'\ \mathit{ind}_i x)] \Rightarrow A\ \mathit{ind}_i \varphi$

Some Properties

- $[(A \textit{ holds}) \textit{ entails } (A' \textit{ holds})] \Rightarrow A \textit{ ind}_i(A' \textit{ holds})$
- $[(A \textit{ ind}_i\ \varphi) \wedge (A \textit{ ind}_i \psi)] \Rightarrow A \textit{ ind}_i(\varphi \wedge \psi)$
- $[(A \textit{ ind}_i[A' \textit{ holds}]) \wedge (A' \textit{ ind}_i x)] \Rightarrow A \textit{ ind}_i \varphi$
- $[(A \textit{ ind}_i \varphi) \wedge (\varphi \textit{ entails } \psi)] \Rightarrow A \textit{ ind}_i \psi$

Some Properties

- $[(A \textit{ holds}) \textit{ entails } (A' \textit{ holds})] \Rightarrow A \textit{ ind}_i(A' \textit{ holds})$
- $[(A \textit{ ind}_i\ \varphi) \wedge (A \textit{ ind}_i \psi)] \Rightarrow A \textit{ ind}_i(\varphi \wedge \psi)$
- $[(A \textit{ ind}_i[A' \textit{ holds}]) \wedge (A' \textit{ ind}_i x)] \Rightarrow A \textit{ ind}_i \varphi$
- $[(A \textit{ ind}_i \varphi) \wedge (\varphi \textit{ entails } \psi)] \Rightarrow A \textit{ ind}_i \psi$
- $[(A \textit{ ind}_i\ R_j[A' \textit{ holds}]) \wedge R_i(A' \textit{ ind}_j \varphi)] \Rightarrow A \textit{ ind}_i R_j(\varphi)$

Reflexive Common Indicator

Reflexive Common Indicator

- A *holds* $\Rightarrow R_i(A$ *holds*$)$

Reflexive Common Indicator

- $A\ holds \Rightarrow R_i(A\ holds)$
- $A\ ind_i\ R_j(A\ holds)$

Reflexive Common Indicator

- $A \text{ holds} \Rightarrow R_i(A \text{ holds})$
- $A \; ind_i \; R_j(A \text{ holds})$
- $A \; ind_i \; \varphi$

Reflexive Common Indicator

- $A\ holds \Rightarrow R_i(A\ holds)$
- $A\ ind_i\ R_j(A\ holds)$
- $A\ ind_i\ \varphi$
- $(A\ ind_i\ \psi) \Rightarrow R_i[A\ ind_j\ \psi]$

Let $R^G(\varphi)$: $R_i\varphi, R_j\varphi, \ldots$, $R_i(R_j\varphi)$, $R_j(R_i(\varphi))$, $\ldots$
iterated reason to believe φ.

Let $R^G(\varphi)$: $R_i\varphi, R_j\varphi, \ldots$, $R_i(R_j\varphi)$, $R_j(R_i(\varphi))$, ...
iterated reason to believe φ.

Theorem. (Lewis) For all states of affairs A, for all propositions φ, and for all groups G: if A holds, and if A is a reflexive common indicator in G that φ, then $R^G(\varphi)$ is true.

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge")

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Example. Suppose there is an agent $i \notin G$ that is *authoritative* for each member of G.

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Example. Suppose there is an agent $i \notin G$ that is *authoritative* for each member of G. So, for $j \in G$, "i states to j that φ is true" *indicates to j that φ*.

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Example. Suppose there is an agent $i \not\in G$ that is *authoritative* for each member of G. So, for $j \in G$, "i states to j that φ is true" *indicates to j that φ*. Suppose that *separately and privately* to each member of G, i states that φ and $R^G(\varphi)$ are true.

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Example. Suppose there is an agent $i \not\in G$ that is *authoritative* for each member of G. So, for $j \in G$, "i states to j that φ is true" *indicates to j that φ*. Suppose that *separately and privately* to each member of G, i states that φ and $R^G(\varphi)$ are true.Then, we have $R^i\varphi$ and $R_i(R^G(\varphi))$ for each $i \in G$.

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Example. Suppose there is an agent $i \notin G$ that is *authoritative* for each member of G. So, for $j \in G$, "i states to j that φ is true" *indicates to* j *that* φ. Suppose that *separately and privately* to each member of G, i states that φ and $R^G(\varphi)$ are true.Then, we have $R^i\varphi$ and $R_i(R^G(\varphi))$ for each $i \in G$. *But there is no common indicator that* φ *is true.*

Lewis and Aumann

Lewis common knowledge that φ *implies* the iterated definition of common knowledge ("Aumann common knowledge"), but the converse is not generally true....

Example. Suppose there is an agent $i \notin G$ that is *authoritative* for each member of G. So, for $j \in G$, "i states to j that φ is true" *indicates to j that φ*. Suppose that *separately and privately* to each member of G, i states that φ and $R^G(\varphi)$ are true.Then, we have $R^i\varphi$ and $R_i(R^G(\varphi))$ for each $i \in G$. *But there is no common indicator that φ is true.* The agents $j \in G$ may have no reason to believe that everyone heard the statement from i or that all agents in G treat i as authoritative.

How does this help?

			B	
		l	r	
A	l	10,10	0,0	
	r	0,0	11,11	

A: What should *we* do? **Team Reasoning**: why should this "mode of reasoning" be endorsed?

Reason to Believe Logic

$R_i(\varphi)$: "agent i has reason to believe φ"

Reason to Believe Logic

$R_i(\varphi)$: "agent i has reason to believe φ" this is interpreted as φ follows from rules (deductive, inductive, norm of practical reason) *endorsed by agent i*.

Reason to Believe Logic

$R_i(\varphi)$: "agent i has reason to believe φ" this is interpreted as φ follows from rules (deductive, inductive, norm of practical reason) *endorsed by agent i*.

Inference rules associated with the Reason-to-believe logic:
$inf(R) : \varphi, \psi \rightarrow \chi$

Reason to Believe Logic

$R_i(\varphi)$: "agent *i* has reason to believe φ" this is interpreted as φ follows from rules (deductive, inductive, norm of practical reason) *endorsed by agent i*.

Inference rules associated with the Reason-to-believe logic:
$\mathit{inf}(R) : \varphi, \psi \to \chi$

Assume each person's logic at least contains propositional logic:
$\mathit{inf}(R) : \varphi_1, \dots \varphi_n, \neg(\varphi_1 \wedge \cdots \wedge \varphi_n \wedge \neg\psi) \to \psi$

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will:

- a prediction about what i will choose in a future decision problem;
- a deontic statement about what i ought to choose;
- assert that i endorses some inference rule; or
- assert that i has reason to believe some proposition

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will:

- a prediction about what i will choose in a future decision problem;
- a deontic statement about what i ought to choose;
- assert that i endorses some inference rule; or
- assert that i has reason to believe some proposition

$R_i(\varphi_i)$ vs. $R_j(\varphi_i)$: Suppose i reliable takes a bus every Monday.

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will:

- a prediction about what i will choose in a future decision problem;
- a deontic statement about what i ought to choose;
- assert that i endorses some inference rule; or
- assert that i has reason to believe some proposition

$R_i(\varphi_i)$ vs. $R_j(\varphi_i)$: Suppose i reliable takes a bus every Monday. The other commuters may all make the inductive inference that i will take the bus next Monday (M_i).

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will:

- a prediction about what i will choose in a future decision problem;
- a deontic statement about what i ought to choose;
- assert that i endorses some inference rule; or
- assert that i has reason to believe some proposition

$R_i(\varphi_i)$ vs. $R_j(\varphi_i)$: Suppose i reliable takes a bus every Monday. The other commuters may all make the inductive inference that i will take the bus next Monday (M_i). In fact, we may assume that this is a *common mode of reasoning*, so everyone reliably makes the inference that i will catch the bus next Monday.

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will:

- a prediction about what i will choose in a future decision problem;
- a deontic statement about what i ought to choose;
- assert that i endorses some inference rule; or
- assert that i has reason to believe some proposition

$R_i(\varphi_i)$ vs. $R_j(\varphi_i)$: Suppose i reliable takes a bus every Monday. The other commuters may all make the inductive inference that i will take the bus next Monday (M_i). In fact, we may assume that this is a *common mode of reasoning*, so everyone reliably makes the inference that i will catch the bus next Monday. So, $R_j(M_i)$, $R_iR_j(M_i)$

Subject of the Proposition

Agent i is the **subject of the proposition** φ_i if φ_i makes an assertion about a current or future act of is will:

- a prediction about what i will choose in a future decision problem;
- a deontic statement about what i ought to choose;
- assert that i endorses some inference rule; or
- assert that i has reason to believe some proposition

$R_i(\varphi_i)$ vs. $R_j(\varphi_i)$: Suppose i reliable takes a bus every Monday. The other commuters may all make the inductive inference that i will take the bus next Monday (M_i). In fact, we may assume that this is a *common mode of reasoning*, so everyone reliably makes the inference that i will catch the bus next Monday. So, $R_j(M_i)$, $R_iR_j(M_i)$, but i should still be *free* to choose whether he wants to take the bus on Monday, so $\neg R_i(M_i)$ and $\neg R_j(R_i(M_i))$, etc.

Common Reason to Believe

Awareness of Common Reason: for all $i \in G$ and all propositions φ,

$$R^G(\varphi) \Rightarrow R_i[R^G(\varphi)]$$

Common Reason to Believe

Awareness of Common Reason: for all $i \in G$ and all propositions φ,

$$R^G(\varphi) \Rightarrow R_i[R^G(\varphi)]$$

Authority of Common Reason: for all $i \in G$ and all propositions φ for which i is not the subject

$$inf(R_i) : R^G(\varphi) \rightarrow \varphi$$

Common Reason to Believe

Awareness of Common Reason: for all $i \in G$ and all propositions φ,

$$R^G(\varphi) \Rightarrow R_i[R^G(\varphi)]$$

Authority of Common Reason: for all $i \in G$ and all propositions φ for which i is not the subject

$$\mathit{inf}(R_i) : R^G(\varphi) \rightarrow \varphi$$

Common Attribution of Common Reason: for all $i \in G$, for all propositions φ for which i is not the subject

$$\mathit{inf}(R^G) : \varphi \rightarrow R_i(\varphi)$$

Common Reason to Believe to Common Belief

Theorem The three previous properties can generate any hierarchy of belief (i has reason to believe that j has reason to believe that... that φ) for any φ with $R^G(\varphi)$.

Team Maximising

$inf(R_i) : R^N[opt(v, N, s^N)]$,
$R^N[$ each $i \in N$ endorses team maximising with respect to N and v $]$,
$R^N[$ each member of N acts on reasons $] \rightarrow ought(i, s_i)$

Team Maximising

$inf(R_i) : R^N[opt(v, N, s^N)],$
$R^N[$ each $i \in N$ endorses team maximising with respect to N and v $],$
$R^N[$ each member of N acts on reasons $] \rightarrow ought(i, s_i)$

$R_i[ought(i, s_i)]$: i has reason to choose s_i

Team Maximising

$inf(R_i) : R^N[opt(v, N, s^N)],$
$R^N[$ each $i \in N$ endorses team maximising with respect to N and v $],$
$R^N[$ each member of N acts on reasons $] \rightarrow ought(i, s_i)$

i acts on reasons if for all s_i, $R_i[ought(i, s_i)] \Rightarrow choice(i, s_i)$

Team Maximising

$inf(R_i) : R^N[opt(v, N, s^N)]$,
$R^N[$ each $i \in N$ endorses team maximising with respect to N and v $]$,
$R^N[$ each member of N acts on reasons $] \rightarrow ought(i, s_i)$

$opt(v, N, s^N)$: s^N is maximal for the group N w.r.t. v

Team Maximising

$inf(R_i) : R^N[opt(v, N, s^N)]$,
$R^N[$ each $i \in N$ endorses team maximising with respect to N and v $]$,
$R^N[$ each member of N acts on reasons $] \rightarrow ought(i, s_i)$

Recursive definition: i's endorsement of the rule depends on i having a reason to believe everyone else endorses the rule...

Different contexts of agency

Different contexts of agency

- Individual decision making and individual action against nature.
  - Ex: Gambling.

Different contexts of agency

- Individual decision making and individual action against nature.

- Individual decision making in interaction.
  - Ex: Playing chess.

Different contexts of agency

- Individual decision making and individual action against nature.
- Individual decision making in interaction.
- Collective decision making.
  - Ex: Carrying the piano.

Different contexts of agency

- Individual decision making and individual action against nature.

- Individual decision making in interaction.

- Collective decision making.


Next: Social Choice Theory and Group Preferences