Epistemic Game Theory

Lecture 10

Eric Pacuit

University of Maryland, College Park
`pacuit.org`
`epacuit@umd.edu`

April 15, 2014

Three Assumptions

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$

Three Assumptions

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$

- A game is a *partial* description of a set (or sequence) of interdependent Bayesian decision problems. The qualitative constraints will not normally be enough to determine what the agents believe about each other, or to determine what solutions are prescribed to the decision problems.

Three Assumptions

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$

- A game is a *partial* description of a set (or sequence) of interdependent Bayesian decision problems. The qualitative constraints will not normally be enough to determine what the agents believe about each other, or to determine what solutions are prescribed to the decision problems.
- There is no concept of rationality for decision making in a situation where the outcomes depend on the actions of more than one agent. The utilities of the other agents are only relevant to an agent as information that, together with beliefs about the rationality of those agents, help predict their actions.

Three Assumptions

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$

- A game is a *partial* description of a set (or sequence) of interdependent Bayesian decision problems. The qualitative constraints will not normally be enough to determine what the agents believe about each other, or to determine what solutions are prescribed to the decision problems.
- There is no concept of rationality for decision making in a situation where the outcomes depend on the actions of more than one agent. The utilities of the other agents are only relevant to an agent as information that, together with beliefs about the rationality of those agents, help predict their actions.
- There are no special rules of rationality telling one what to do in the absence of degrees of belief except: decide what you believe, and then maximize expected utility.

Models of Games

Suppose that G is a game.

- Outcomes of the game: $S = \Pi_{i \in N} S_i$
- The players' beliefs, or conjectures: $\{P_i\}_{i \in N}$, $P_i \in \Delta(S_{-i})$

Models of Games, continued

A model of a game $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$ is a completion of the partial specification of the Bayesian decision problems and a representation of a particular play of the game.

Models of Games, continued

A model of a game $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$ is a completion of the partial specification of the Bayesian decision problems and a representation of a particular play of the game.

- W os a set of *possible worlds*
- $\mathbf{s}$ is a function $\mathbf{s} : W \to \Pi_{i\in N} S_i$

Models of Games, continued

A model of a game $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$ is a completion of the partial specification of the Bayesian decision problems and a representation of a particular play of the game.

- W os a set of *possible worlds*
- $\mathbf{s}$ is a function $\mathbf{s} : W \to \Pi_{i \in N} S_i$
- For $s_i \in S_i$, $[s] = \{w \mid \mathbf{s}(w) = s\}$, if $X \subseteq S$, $[X] = \bigcup_{s \in X} [s]$, $[s_i] = \{w \mid \mathbf{s}_i(w) = s_i\}$.

Models of Games, continued

A model of a game $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$ is a completion of the partial specification of the Bayesian decision problems and a representation of a particular play of the game.

- W os a set of *possible worlds*
- $\mathbf{s}$ is a function $\mathbf{s} : W \rightarrow \Pi_{i\in N} S_i$
- For $s_i \in S_i$, $[s] = \{w \mid \mathbf{s}(w) = s\}$, if $X \subseteq S$, $[X] = \bigcup_{s\in X}[s]$, $[s_i] = \{w \mid \mathbf{s}_i(w) = s_i\}$.
- For each $i \in N$, $P_i \in \Delta(W)$. Two assumptions:
  - $[s]$ is measurable for all strategy profiles $s \in S$
  - $P_i([s_i]) > 0$ for all $s_i \in S_i$

Models of Games, continued

A model of a game $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$ is a completion of the partial specification of the Bayesian decision problems and a representation of a particular play of the game.

- W os a set of *possible worlds*
- $\mathbf{s}$ is a function $\mathbf{s} : W \to \Pi_{i \in N} S_i$
- For $s_i \in S_i$, $[s] = \{w \mid \mathbf{s}(w) = s\}$, if $X \subseteq S$, $[X] = \bigcup_{s \in X}[s]$, $[s_i] = \{w \mid \mathbf{s}_i(w) = s_i\}$.
- For each $i \in N$, $P_i \in \Delta(W)$. Two assumptions:
  - $[s]$ is measurable for all strategy profiles $s \in S$
  - $P_i([s_i]) > 0$ for all $s_i \in S_i$
- $P_{i,w}(E) = P_i(E \mid [\sigma_i(w)]) = \frac{P_i(E \cap [\sigma_i(w))}{P_i([\sigma_i(w)])}$

Models of Games, continued

A model of a game $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$ is a completion of the partial specification of the Bayesian decision problems and a representation of a particular play of the game.

- W os a set of *possible worlds*
- $\mathbf{s}$ is a function $\mathbf{s} : W \rightarrow \Pi_{i\in N} S_i$
- For $s_i \in S_i$, $[s] = \{w \mid \mathbf{s}(w) = s\}$, if $X \subseteq S$, $[X] = \bigcup_{s\in X}[s]$, $[s_i] = \{w \mid \mathbf{s}_i(w) = s_i\}$.
- For each $i \in N$, $P_i \in \Delta(W)$. Two assumptions:
  - $[s]$ is measurable for all strategy profiles $s \in S$
  - $P_i([s_i]) > 0$ for all $s_i \in S_i$
- $P_{i,w}(E) = P_i(E \mid [\sigma_i(w)]) = \frac{P_i(E\cap[\sigma_i(w))}{P_i([\sigma_i(w)])}$

$$P_i([s]) \text{ vs. } P_{i,w}([s])$$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_1, s_{-i})$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_1, s_{-i})$

$\text{Rat}_i = \{w \mid EU_{i,w}(\mathbf{s}_i(w)) \geq EU_{i,w}(s_i) \text{ for all } s_i \in S_i\}$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_1, s_{-i})$

$\mathsf{Rat}_i = \{w \mid EU_{i,w}(\mathbf{s}_i(w)) \geq EU_{i,w}(s_i) \text{ for all } s_i \in S_i\}$

Each $P \in \Delta(W)$ is associated with $P^S \in \Delta(S)$ as follows: for all $s \in S$, $P^S(s) = P([s])$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_1, s_{-i})$

$\mathrm{Rat}_i = \{w \mid EU_{i,w}(\mathbf{s}_i(w)) \geq EU_{i,w}(s_i) \text{ for all } s_i \in S_i\}$

Each $P \in \Delta(W)$ is associated with $P^S \in \Delta(S)$ as follows: for all $s \in S$, $P^S(s) = P([s])$

A mixed strategy $\sigma \in \Pi_{i \in N} \Delta(S_i)$, $P_\sigma \in \Delta(S)$, $P_\sigma(s) = \sigma_1(s_1) \cdots \sigma_n(s_n)$

Theorem (Aumann). σ is a Nash equilibrium of G iff there exists a model $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$ such that for all $i \in N$, $\mathrm{Rat}_i = W$, for all $i, j \in N$, $P_i = P_j$ and for all $i \in N$, $P_i^S = P_\sigma$.

Theorem (Aumann). σ is a correlated equilibrium of G iff there exists a model $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$ such that for all $i \in N$, $\mathrm{Rat}_i = W$ and for all $i \in N$, $P_i^S = \sigma$.

A **best reply set** (BRS) is a sequence $(B_1, B_2, \ldots, B_n) \subseteq S = \Pi_{i \in N} S_i$ such that for all $i \in N$, and all $s_i \in B_i$, there exists $\mu_{-i} \in \Delta(B_{-i})$ such that s_i is a best response to μ_{-i}: I.e.,

$$b_i = \arg\max_{s_i \in S_i} EU_{i,\mu_{-i}}(s_i)$$

.

		2			
		b_1	b_2	b_3	b_4
1	a_1	0, 7	2, 5	7, 0	0, 1
	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

2

		b_1	b_2	b_3	b_4
	a_1	0, 7	2, 5	7, 0	0, 1
1	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

- (a_2, b_2) is the unique Nash equilibria, hence $(\{a_2\}, \{b_2\})$ is a BRS

		2			
		b_1	b_2	b_3	b_4
1	a_1	0, 7	2, 5	7, 0	0, 1
	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

- (a_2, b_2) is the unique Nash equilibria, hence $(\{a_2\}, \{b_2\})$ is a BRS
- $(\{a_1, a_3\}, \{b_1, b_3\})$ is a BRS

2

		b_1	b_2	b_3	b_4
	a_1	0, 7	2, 5	7, 0	0, 1
1	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

- (a_2, b_2) is the unique Nash equilibria, hence $(\{a_2\}, \{b_2\})$ is a BRS
- $(\{a_1, a_3\}, \{b_1, b_3\})$ is a BRS
- $(\{a_1, a_2, a_3\}, \{b_1, b_2, b_3\})$ is a full BRS

Theorem (Bernheim; Pearce; Brandenburger and Dekel; …). $(B_1, B_2, \ldots, B_n)$ is a BRS for G iff there exists a model $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$ such that for all $i \in N$, $\mathsf{Rat}_i = W$ and $[B_1 \times \cdots \times B_n] = W$.

Epistemic and Causal Possibilities

"In deliberation, I reason both about how the world might have been different if I or others did different things than we are going to do, and also about how my beliefs, or others' beliefs, might change if I or they learned things that we expect not to learn." (pg. 134)

R. Stalnaker. *Knowledge, Belief and Counterfactual Reasoning in Games*. Economics and Philosophy, 12, pgs. 133 - 163, 1996.

A game model $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$ can represent the first type of possibility but not the second.

In a game model $\mathcal{M}^G$ different states represent different beliefs only when the agent is doing something different.

To represent different beliefs, we need a set of models $\{\mathcal{M}_1^G, \mathcal{M}_2^G, \ldots\}$.

In a game model $\mathcal{M}^G$ different states represent different beliefs only when the agent is doing something different.

To represent different beliefs, we need a set of models $\{\mathcal{M}_1^G, \mathcal{M}_2^G, \ldots\}$.

Alternatively, $\mathcal{M}^G = \langle W, \{R_i, P_i\}_{i \in N}, \mathbf{s} \rangle$. Where W and **s** are as before. $w \; R_i \; v$ means v is compatible with what i believes in w.

In a game model $\mathcal{M}^G$ different states represent different beliefs only when the agent is doing something different.

To represent different beliefs, we need a set of models $\{\mathcal{M}_1^G, \mathcal{M}_2^G, \ldots\}$.

Alternatively, $\mathcal{M}^G = \langle W, \{R_i, P_i\}_{i \in N}, \mathbf{s} \rangle$. Where W and $\mathbf{s}$ are as before. $w \; R_i \; v$ means v is compatible with what i believes in w.

- Then, the set $R_i(w) = \{v \mid w \; R_i \; v\}$ is the set of all worlds compatible with what i believes in w

In a game model $\mathcal{M}^G$ different states represent different beliefs only when the agent is doing something different.

To represent different beliefs, we need a set of models $\{\mathcal{M}^G_1, \mathcal{M}^G_2, \ldots\}$.

Alternatively, $\mathcal{M}^G = \langle W, \{R_i, P_i\}_{i \in N}, \mathbf{s}\rangle$. Where W and $\mathbf{s}$ are as before. $w\ R_i\ v$ means v is compatible with what i believes in w.

- Then, the set $R_i(w) = \{v \mid w\ R_i\ v\}$ is the set of all worlds compatible with what i believes in w
- Assume that R_i is serial, transitive and Euclidean. So, for any w, v if the agent has different beliefs in w and v, then $R_i(w) \cap R_i(v) = \emptyset$.

In a game model $\mathcal{M}^G$ different states represent different beliefs only when the agent is doing something different.

To represent different beliefs, we need a set of models $\{\mathcal{M}^G_1, \mathcal{M}^G_2, \ldots\}$.

Alternatively, $\mathcal{M}^G = \langle W, \{R_i, P_i\}_{i \in N}, \mathbf{s} \rangle$. Where W and $\mathbf{s}$ are as before. $w\ R_i\ v$ means v is compatible with what i believes in w.

- Then, the set $R_i(w) = \{v \mid w\ R_i\ v\}$ is the set of all worlds compatible with what i believes in w
- Assume that R_i is serial, transitive and Euclidean. So, for any w, v if the agent has different beliefs in w and v, then $R_i(w) \cap R_i(v) = \emptyset$.
- Assume that for all w, v, if $w\ R_i\ v$, then $\mathbf{s}_i(v) = \mathbf{s}_i(w)$

In a game model $\mathcal{M}^G$ different states represent different beliefs only when the agent is doing something different.

To represent different beliefs, we need a set of models $\{\mathcal{M}^G_1, \mathcal{M}^G_2, \ldots\}$.

Alternatively, $\mathcal{M}^G = \langle W, \{R_i, P_i\}_{i \in N}, \mathbf{s}\rangle$. Where W and $\mathbf{s}$ are as before. $w\ R_i\ v$ means v is compatible with what i believes in w.

- Then, the set $R_i(w) = \{v \mid w\ R_i\ v\}$ is the set of all worlds compatible with what i believes in w
- Assume that R_i is serial, transitive and Euclidean. So, for any w, v if the agent has different beliefs in w and v, then $R_i(w) \cap R_i(v) = \emptyset$.
- Assume that for all w, v, if $w\ R_i\ v$, then $\mathbf{s}_i(v) = \mathbf{s}_i(w)$
- $P_{i,w}(E) = P_i(E \mid R_i(w))$
- Rat_i is defined as before

Enough Counterfactual Possibilities

If a player had chosen a different strategy from the one he in fact chose, the other players would still have chosen the same strategies, and would have had the same beliefs, that they in fact had.

Enough Counterfactual Possibilities

If a player had chosen a different strategy from the one he in fact chose, the other players would still have chosen the same strategies, and would have had the same beliefs, that they in fact had.

For any world w and strategy $s_i \in S_i$ for player i, there is a world $f(w, s)$ such that

1. for all $j \neq i$, if $w\ R_j\ v$, then $f(w, s_i)\ R_j\ v$
2. if $w\ R_i\ v$, then $f(w, s_i)\ R_i\ f(v, s_i)$
3. $\mathbf{s}_i(f(w, s_i)) = s_i$
4. $P_i(f(w, s)) = P_i(w)$

Enough Counterfactual Possibilities

If a player had chosen a different strategy from the one he in fact chose, the other players would still have chosen the same strategies, and would have had the same beliefs, that they in fact had.

For any world w and strategy $s_i \in S_i$ for player i, there is a world $f(w, s)$ such that

1. for all $j \neq i$, if $w\ R_j\ v$, then $f(w, s_i)\ R_j\ v$
2. if $w\ R_i\ v$, then $f(w, s_i)\ R_i\ f(v, s_i)$
3. $\mathbf{s}_i(f(w, s_i)) = s_i$
4. $P_i(f(w, s)) = P_i(w)$

$f(w, s_i)$ represents the counterfactual possible world that, in w, is the world that would have been realized if player i, believing exactly what he believes in w about the other players, had chosen strategy s_i.

"Even if I am resolved to act rationally, I may consider in deliberation what the consequences would be of acting in ways that are not. And even if I am certain that you will act rationally, I may consider how I would revise my beliefs if I learn that I was wrong about this."

(pg. 14, Stalnaker)

As is standard, we suppose that the agent revises her beliefs by conditionalization, but nothing in the models describes how the players revise her beliefs if they learn something that had a prior probability 0.

As is standard, we suppose that the agent revises her beliefs by conditionalization, but nothing in the models describes how the players revise her beliefs if they learn something that had a prior probability 0.

"...even in a static situation, one might ask how an agent's beliefs are disposed to change were he to learn that he was mistaken about something he believe with probability one, and the answer to this question may be relevant to his decisions." (pg. 143, Stalnaker)

Let $P \in \Delta(X)$ be a probability measure, the **support** of P is $\mathrm{supp}(P) = \{x \in X \mid P(x) > 0\}$.

A probability measure $P \in \Delta(X)$ is said to be a **full support** probability measure on X provided $\mathrm{supp}(P) = X$.

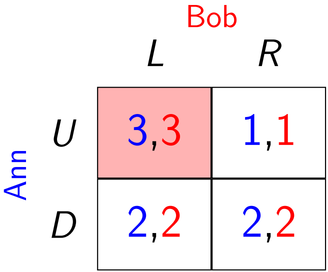
		Bob	
		L	R
Ann	U	3,3	1,1
	D	2,2	2,2

Is D rationalizable?

		Bob	
		L	*R*
Ann	*U*	3,3	1,1
	D	2,2	2,2

Is *D* rationalizable?

There is no *full support* probability such that *R* is a best response



		Bob	
		L	R
Ann	U	3,3	1,1
	D	2,2	2,2

Is D rationalizable?

There is no *full support* probability such that R is a best response

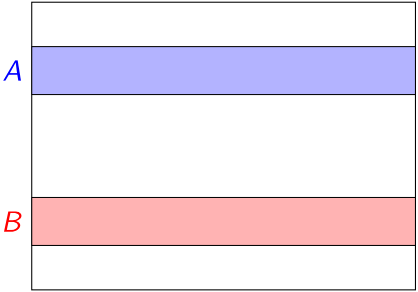
Should Ann assign probability 0 to R or probability > 0 to R?

Strategic Reasoning and Admissibility

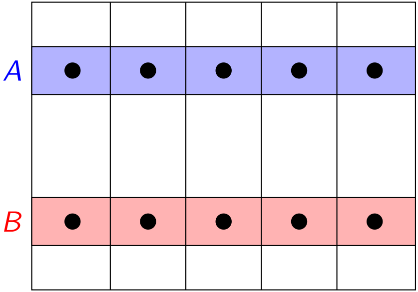
"The argument for deletion of a weakly dominated strategy for player *i* is that he contemplates the possibility that every strategy combination of his rivals occurs with positive probability. However, this hypothesis clashes with the logic of iterated deletion, which assumes, precisely, that eliminated strategies are not expected to occur."

Mas-Colell, Whinston and Green. *Introduction to Microeconomics*. 1995.

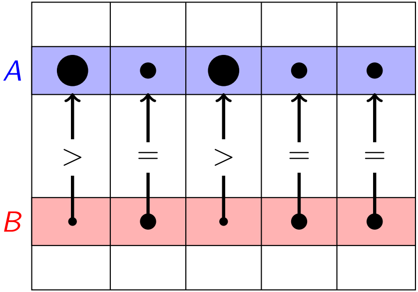
Weak Dominance



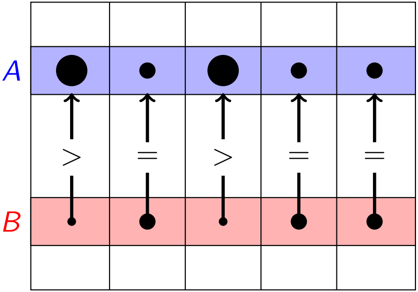
Weak Dominance



Weak Dominance



Weak Dominance



Privacy of tie-breaking: If a strategy a is optimal for player j, then player i cannot *know* that j will not choose a.

A Puzzle

R. Cubitt and R. Sugden. *Rationally Justifiable Play and the Theory of Non-cooperative games*. Economic Journal, 104, pgs. 798 - 803, 1994.

R. Cubitt and R. Sugden. *Common reasoning in games: A Lewisian analysis of common knowledge of rationality*. Manuscript, 2011.

A Puzzle

A Puzzle

	in_2	out_2
in_1	$1, 1, 1$	$1, 1, 1$
out_1	$1, 1, 1$	$0, 1, 1$

in_3

	in_2	out_2
in_1	$1, 1, 1$	$1, 0, 1$
out_1	$1, 1, 0$	$0, 0, 0$

out_3

A Puzzle

	in_2	out_2
in_1	1, 1, 1	1, 1, 1
out_1	1, 1, 1	0, 1, 1

in_3

	in_2	out_2
in_1	1, 1, 1	1, 0, 1
out_1	1, 1, 0	0, 0, 0

out_3

1. If 1 considers out_2 possible, then it is common knowledge that out_1 is not possible

A Puzzle

	in_2	out_2
in_1	$1, 1, 1$	$1, 1, 1$
out_1	$1, 1, 1$	$0, 1, 1$

in_3

	in_2	out_2
in_1	$1, 1, 1$	$1, 0, 1$
out_1	$1, 1, 0$	$0, 0, 0$

out_3

1. If 1 considers out_2 possible, then it is common knowledge that out_1 is not possible
2. If 2 considers out_3 possible, then it is common knowledge that out_2 is not possible

A Puzzle

	in_2	out_2
in_1	$1, 1, 1$	$1, 1, 1$
out_1	$1, 1, 1$	$0, 1, 1$

in_3

	in_2	out_2
in_1	$1, 1, 1$	$1, 0, 1$
out_1	$1, 1, 0$	$0, 0, 0$

out_3

1. If 1 considers out_2 possible, then it is common knowledge that out_1 is not possible
2. If 2 considers out_3 possible, then it is common knowledge that out_2 is not possible
3. If 3 considers out_1 possible, then it is common knowledge that out_3 is not possible

A Puzzle

	in_2	out_2
in_1	$1, 1, 1$	$1, 1, 1$
out_1	$1, 1, 1$	$0, 1, 1$

in_3

	in_2	out_2
in_1	$1, 1, 1$	$1, 0, 1$
out_1	$1, 1, 0$	$0, 0, 0$

out_3

4. If 1 does not consider out_2 possible, then 2 & 3 must consider in_1 & out_1 possible

A Puzzle

	in_2	out_2
in_1	1, 1, 1	1, 1, 1
out_1	1, 1, 1	0, 1, 1

in_3

	in_2	out_2
in_1	1, 1, 1	1, 0, 1
out_1	1, 1, 0	0, 0, 0

out_3

4. If 1 does not consider out_2 possible, then 2 & 3 must consider in_1 & out_1 possible
5. If 2 does not consider out_3 possible, then 1 & 3 must consider in_2 & out_2 possible

A Puzzle

	in_2	out_2
in_1	$1, 1, 1$	$1, 1, 1$
out_1	$1, 1, 1$	$0, 1, 1$

in_3

	in_2	out_2
in_1	$1, 1, 1$	$1, 0, 1$
out_1	$1, 1, 0$	$0, 0, 0$

out_3

4. If 1 does not consider out_2 possible, then 2 & 3 must consider in_1 & out_1 possible
5. If 2 does not consider out_3 possible, then 1 & 3 must consider in_2 & out_2 possible
6. If 3 does not consider out_1 possible, then 1 & 2 must consider in_3 & out_3 possible

A Puzzle

	in_2	out_2
in_1	1, 1, 1	1, 1, 1
out_1	1, 1, 1	0, 1, 1

in_3

	in_2	out_2
in_1	1, 1, 1	1, 0, 1
out_1	1, 1, 0	0, 0, 0

out_3

- If i considers out_{i+1} possible, then it is common knowledge that out_i is not possible
- If i does not consider out_{i+1} possible, then $i+1$ & $i+2$ must consider in_i & out_i possible

A Puzzle

	in_2	out_2
in_1	$1, 1, 1$	$1, 1, 1$
out_1	$1, 1, 1$	$0, 1, 1$

in_3

	in_2	out_2
in_1	$1, 1, 1$	$1, 0, 1$
out_1	$1, 1, 0$	$0, 0, 0$

out_3

- If i considers out_{i+1} possible, then it is common knowledge that out_i is not possible
- If i does not consider out_{i+1} possible, then $i + 1$ & $i + 2$ must consider in_i & out_i possible
- 1 does consider out_2 possible $\implies$ 3 does not consider out_1 possible $\implies$ 2 considers out_3 possible $\implies$ 1 does not consider out_2 possible

Let $G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$ be a game.

A strategy is *justifiable* if and only if it is optimal with respect to some *coherent* set of beliefs. A set of beliefs is coherent if and only if it is internally consistent and satisfies a principle of *caution*.

Given $P_i \in \Delta(S_{-i})$ and $s_k \in S_k$ the marginal of P_i on s_k is

$$P_i[s_k] = \sum_{s_{-i} \in S_{-i}, (s_{-i})_k = s_k} P_i(s_{-i})$$

A justifiable theory is a set of pairs for each $i \in N$, (J_i, C_i) where

1. For all $i \in N$, $J_i \subseteq S_i$ and $C_i \subseteq \Delta(S_{-i})$
2. For all $i \in N$, $C_i \neq \emptyset$

A justifiable theory is a set of pairs for each $i \in N$, (J_i, C_i) where

1. For all $i \in N$, $J_i \subseteq S_i$ and $C_i \subseteq \Delta(S_{-i})$
2. For all $i \in N$, $C_i \neq \emptyset$

3. For all $i \in N$, $s_i \in S_i$, if there exists some $P_i \in C_i$ such that $s_i = \arg\max_{x \in S_i} EU_{i,P_i}(x)$, then $s_i \in J_i$
4. For all $i \in N$, $s_i \in S_i$, if $s_i \in J_i$, then there exists some $P_i \in C_i$ such that $s_i = \arg\max_{x \in S_i} EU_{i,P_i}(x)$

A justifiable theory is a set of pairs for each $i \in N$, (J_i, C_i) where

1. For all $i \in N$, $J_i \subseteq S_i$ and $C_i \subseteq \Delta(S_{-i})$
2. For all $i \in N$, $C_i \neq \emptyset$

3. For all $i \in N$, $s_i \in S_i$, if there exists some $P_i \in C_i$ such that $s_i = \arg\max_{x \in S_i} EU_{i,P_i}(x)$, then $s_i \in J_i$
4. For all $i \in N$, $s_i \in S_i$, if $s_i \in J_i$, then there exists some $P_i \in C_i$ such that $s_i = \arg\max_{x \in S_i} EU_{i,P_i}(x)$

5. For all $i, k \in N$, $P_i \in \Delta(S_{-i})$, if $s_k \not\in J_k$ and $P_i[s_k] > 0$, then $P_i \not\in C_i$
6. For all $i, k \in N$, $P_i \in \Delta(S_{-i})$, if $s_k \in J_k$ and $P_i[s_k] = 0$, then $P_i \not\in C_i$

There is no justifiable theory for the game G.
Let $h(1) = 2$, $h(2) = 3$, $h(3) = 1$.

1. If $J_i = \{in_i, out_i\}$, then $J_{h(i)} = \{in_{h(i)}\}$
2. If $J_i = \{in_i\}$, then $J_{h(i)} = \{in_{h(i)}, out_{h(i)}\}$

There is no justifiable theory for the game G.
Let $h(1) = 2$, $h(2) = 3$, $h(3) = 1$.

1. If $J_i = \{in_i, out_i\}$, then $J_{h(i)} = \{in_{h(i)}\}$
2. If $J_i = \{in_i\}$, then $J_{h(i)} = \{in_{h(i)}, out_{h(i)}\}$

Three cases:

1. $J_1 = \{in_1\} \Longrightarrow J_2 = \{in_2, out_2\} \Longrightarrow J_3 = \{in_3\} \Longrightarrow J_1 = \{in_1, out_1\}$, contradiction.
2. $J_1 = \{out_1\}$. Then, there is $P_1 \in C_1$ such that $P_1[in_2] = 1$. Then, since $in_1 = \arg\max_{x \in S_1} EU_{1,P_1}(x)$, by 3., $in_1 \in J_1$, contradiction.
3. $J_1 = \{in_1, out_1\} \Longrightarrow J_2 = \{in_2\} \Longrightarrow J_3 = \{in_3, out_3\} \Longrightarrow J_1 = \{in_1\}$, contradiction.

Bob

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

Game 1

Bob

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

Game 2

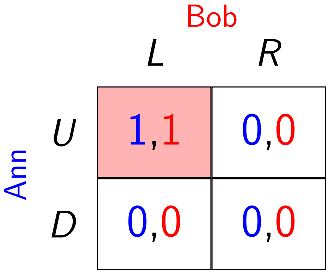
		Bob	
		L	R
Ann	U	1,1	0,0
	D	0,0	0,0

Game 1

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

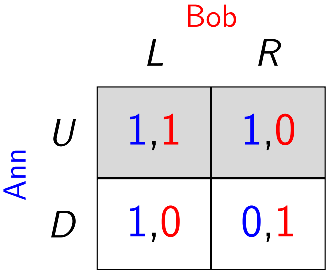
Game 2

Game 1: U weakly dominates D and L weakly dominates R.



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

Game 1



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

Game 2

Game 1: U weakly dominates D and L weakly dominates R.

Game 2: U weakly dominates D

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

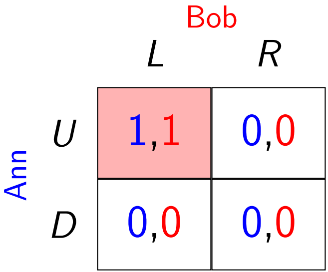
Game 1

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

Game 2

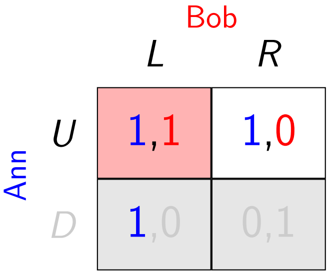
Game 1: U weakly dominates D and L weakly dominates R.

Game 2: U weakly dominates D, and *after removing* D, L strictly dominates R.



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

Game 1

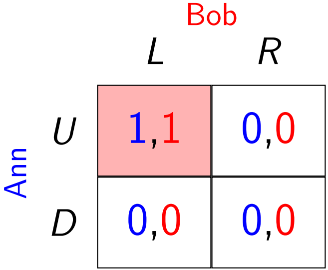


		Bob	
		L	R
Ann	U	1,1	1,0
	D	1,0	0,1

Game 2

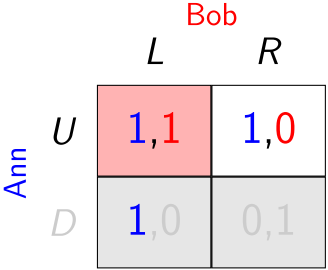
Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

Game 2: But, now what is the reason for not playing *D*?



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

Game 1



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

Game 2

Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

Game 2: But, now what is the reason for not playing *D*?

Theorem (Samuelson). There is no model of Game 2 satisfying common knowledge of rationality (where rationality incorporates weak dominance).

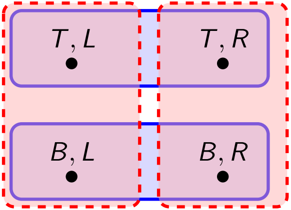
Common Knowledge of Admissibility

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

There is no model of this game with *common knowledge* of admissibility.

Common Knowledge of Admissibility

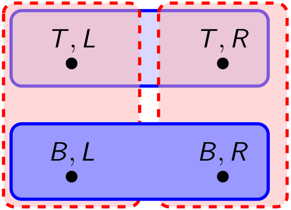
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



The ”full” model of the game

Common Knowledge of Admissibility

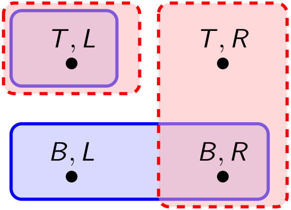
		Bob	
		L	*R*
Ann	*T*	1,1	1,0
	B	1,0	0,1



The "full" model of the game: *B is not admissible given Ann's information*

Common Knowledge of Admissibility

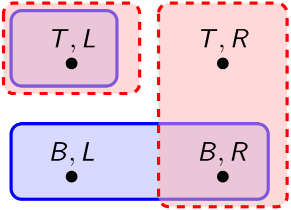
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



What is wrong with this model?

Common Knowledge of Admissibility

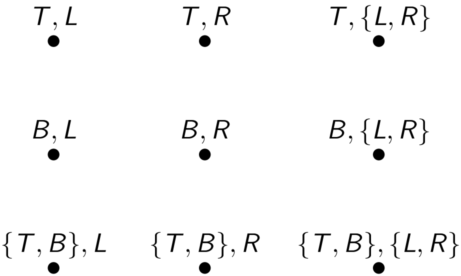
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



Privacy of Tie-Breaking/No Extraneous Beliefs: If a strategy is *rational* for an opponent, then it cannot be "ruled out".

Common Knowledge of Admissibility

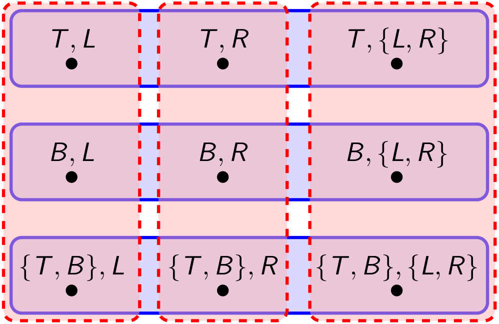
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



Moving to choice *sets*.

Common Knowledge of Admissibility

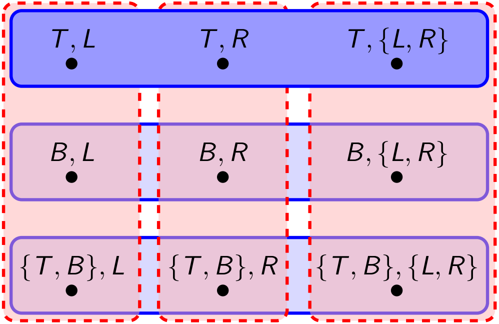
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



Moving to choice *sets*.

Common Knowledge of Admissibility

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



Ann thinks: Bob *has a reason to play* L OR Bob *has a reason to play* R OR Bob *has not yet settled on a choice*

Common Knowledge of Admissibility

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

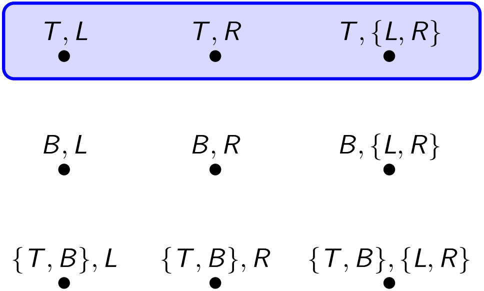
T, L ●	T, R ●	$T, \{L, R\}$ ●
B, L ●	B, R ●	$B, \{L, R\}$ ●
$\{T, B\}, L$ ●	$\{T, B\}, R$ ●	$\{T, B\}, \{L, R\}$ ●

Still there is no model with common knowledge that players have *admissibility-based reasons*

Common Knowledge of Admissibility

Bob

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



T, L T, R $T, \{L, R\}$

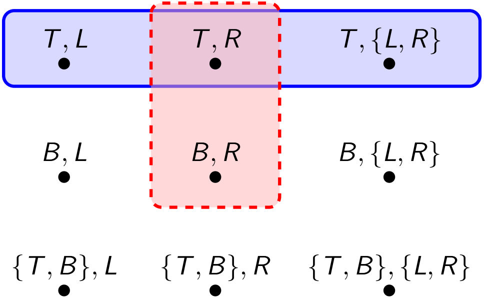
B, L B, R $B, \{L, R\}$

$\{T, B\}, L$ $\{T, B\}, R$ $\{T, B\}, \{L, R\}$

there is a reason to play T provided Ann considers it possible that Bob might play R (actually three cases to consider here)

Common Knowledge of Admissibility

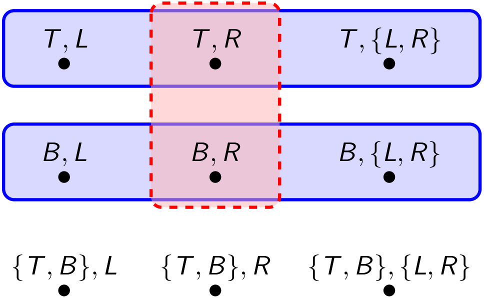
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



But there is a reason to play R provided it is possible that Ann has a reason to play B

Common Knowledge of Admissibility

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



T, L T, R $T, \{L, R\}$

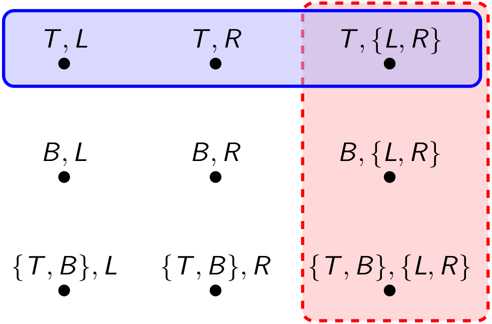
B, L B, R $B, \{L, R\}$

$\{T, B\}, L$ $\{T, B\}, R$ $\{T, B\}, \{L, R\}$

But, there is no reason to play B if there is a reason for Bob to play R.

Common Knowledge of Admissibility

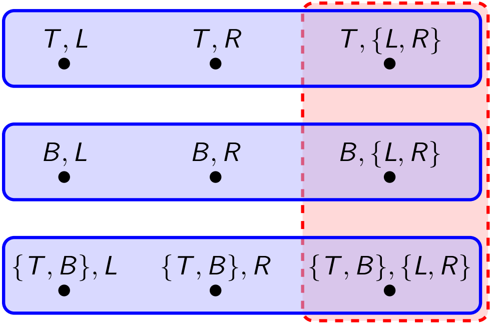
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



R can be ruled out unless there is a possibility that B will be played.

Common Knowledge of Admissibility

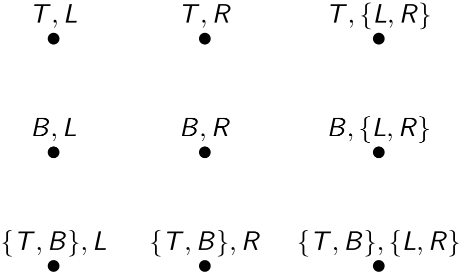
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



there is no reason to play B if R is a *possible play for Bob*.

Common Knowledge of Admissibility

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



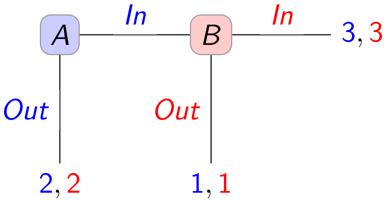
T, L $\quad$ T, R $\quad$ $T, \{L, R\}$

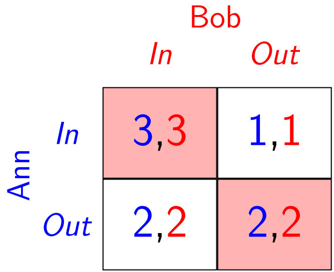
B, L $\quad$ B, R $\quad$ $B, \{L, R\}$

$\{T, B\}, L$ $\quad$ $\{T, B\}, R$ $\quad$ $\{T, B\}, \{L, R\}$

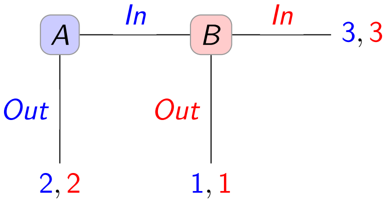
We can check all the possibilities and see we cannot find a model...

		Bob	
		In	*Out*
Ann	*In*	3,3	1,1
	Out	2,2	2,2





		Bob	
		In	*Out*
Ann	*In*	3,3	1,1
	Out	2,2	2,2



- (Out, Out) is rationalizable: Bob assigns probability 1 to Ann choosing *Out* and Bob assigns probability 1 to Ann choosing *Out*. Both are rational given these beliefs.
- What Ann believes Bob will do depends on an event that both Ann and Bob assign probability 0 to.

CPS (Popper Space)

A **conditional probability space** (CPS) over $(W, \mathcal{F})$ is a tuple $(W, \mathcal{F}, \mathcal{F}', \mu)$ such that $\mathcal{F}$ is an algebra over W, $\mathcal{F}'$ is a set of subsets of W (not necessarily an algebra) that does not contain $\emptyset$ and $\mu : \mathcal{F} \times \mathcal{F}' \to [0,1]$ satisfying the following conditions:

1. $\mu(U \mid U) = 1$ if $U \in \mathcal{F}'$
2. $\mu(E_1 \cup E_1 \mid U) = \mu(E_1 \mid U) + \mu(E_2 \mid U)$ if $E_1 \cap E_2 = \emptyset$, $U \in F'$ and $E_1, E_2 \in \mathcal{F}$
3. $\mu(E \mid U) = \mu(E \mid X) \times \mu(X \mid U)$ if $E \subseteq X \subseteq U$, $U, X \in \mathcal{F}'$ and $E \in \mathcal{F}$.

LPS (Lexicographic Probability Space)

A **lexicographic probability space** (LPS) (of length α) is a tuple $(W, \mathcal{F}, \vec{\mu})$ where W is a set of possible worlds, $\mathcal{F}$ is an algebra over W and $\vec{\mu}$ is a sequence of (finitely/countable additive) probability measures on $(W, \mathcal{F})$ indexed by ordinals $< \alpha$.

Suppose that $W = \{w_1, w_2\}$, $\mu_0(w_1) = \mu_0(w_2) = 1/2$ and $\mu_1(w_1) = 1$. The LPS $\vec{\mu} = (\mu_0, \mu_1)$ can be thought of as describing a situation where w_1 is "very slightly" more likely that w_2.

Suppose that X_1 is a bet that pays off 1 if w_1 and 0 in state w_2 and X_2 is a bet that pays off 1 if w_2 and 0 in state w_1.

Then, according to $\vec{\mu}$, X_1 should be "slightly preferred" to a X_2, but for all real numbers $r > 1$, the bet rX_2 is preferred to X_1.

NPS (non-standard probability measures)

$\mathbb{R}^*$ is a *non-Archimedean* field that includes the real numbers as a subfield but also has *infinitesimals*.

For all $b \in \mathbb{R}^*$ such that $-r < b < r$ for some $r \in \mathbb{R}$, there is a unique closest real number a such that $|a - b|$ is an infinitesimal. Let $st(b)$ denote the closest standard real to b.

NPS (non-standard probability measures)

$\mathbb{R}^*$ is a *non-Archimedean* field that includes the real numbers as a subfield but also has *infinitesimals*.

For all $b \in \mathbb{R}^*$ such that $-r < b < r$ for some $r \in \mathbb{R}$, there is a unique closest real number a such that $|a - b|$ is an infinitesimal. Let $st(b)$ denote the closest standard real to b.

A **nonstandard probability space** (NPS) is a tuple $(W, \mathcal{F}, \mu)$ where W is a set of possible worlds, $\mathcal{F}$ is an algebra over W and μ assigns to elements of $\mathcal{F}$, nonnegative elements of $\mathbb{R}^*$ such that $\mu(W) = 1$, $\mu(E \cup F) = \mu(E) + \mu(F)$ if E and F are disjoint.

J. Halpern. *Lexicographic probability, conditional probability, and nonstandard probability*. Games and Economic Behavior, 68:1, pgs. 155 - 179, 2010.

Both Including and Excluding a Strategy

Returning to the problem of weakly dominated strategies and rationalizability, one solution is to assume that players consider some strategies *infinitely more likely than other strategies*.

Both Including and Excluding a Strategy

Returning to the problem of weakly dominated strategies and rationalizability, one solution is to assume that players consider some strategies *infinitely more likely than other strategies*.

		Bob	
		1	[1]
		L	*R*
Ann	*U*	3,3	1,1
	D	2,2	2,2

L. Blume, A. Brandenburger, E. Dekel. *Lexicographic probabilities and choice under uncertainty*. Econometrica, 59(1), pgs. 61 - 79, 1991.

Self-Admissible Sets, I

Let $G = (\{a, b\}, \{S_a, u_a\}, \{S_b, u_b\})$ be a two player game.

If $X \subseteq S_a$ and $Y \subseteq S_b$, $s \in S_a$ is admissible with respect to $X \times Y$ for a if and only if there is a probability measure $P \in \Delta(S_b)$ such that

$$s_a = \arg\max_{x \in X} EU_{a,P}(x)$$

and $\text{supp}(P) = Y$.

Self-Admissible Sets, I

Let $G = (\{a, b\}, \{S_a, u_a\}, \{S_b, u_b\})$ be a two player game.

If $X \subseteq S_a$ and $Y \subseteq S_b$, $s \in S_a$ is admissible with respect to $X \times Y$ for a if and only if there is a probability measure $P \in \Delta(S_b)$ such that

$$s_a = \arg\max_{x \in X} EU_{a,P}(x)$$

and $\text{supp}(P) = Y$.

Fix $s_a \in S_a$, a **strategy** $r_a \in Q_a$ **supports** s_a if and only if there is a $\sigma \in \Delta(S_a)$ such that $U_a(\sigma, s_b) = U_a(s_a, s_b)$ and $r_a \in \text{supp}(\sigma)$.

Self-Admissible Sets, II

A **self-admissible set** is a pair (Q_a, Q_b) such that

1. Each $s_a \in S_a$ is admissible with respect to $S_a \times S_b$
2. Each $s_a \in Q_a$ is admissible with respect to $S_a \times Q_b$
3. For all $s_a \in Q_a$, if r_a *supports* s_a, then $r_a \in Q_a$

SAS Example

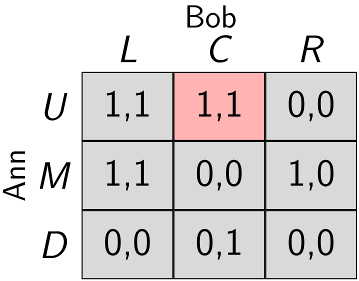
			Bob	
		L	*C*	*R*
	U	1,1	1,1	0,0
Ann	*M*	1,1	0,0	1,0
	D	0,0	0,1	0,0

SAS Example

			Bob	
		L	C	R
Ann	U	1,1	1,1	0,0
	M	1,1	0,0	1,0
	D	0,0	0,1	0,0

Five SASs: $\{(U, L)\}$,

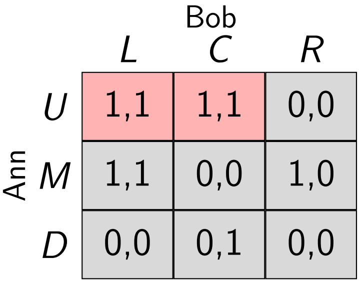
SAS Example



			Bob	
		L	C	R
	U	1,1	1,1	0,0
Ann	M	1,1	0,0	1,0
	D	0,0	0,1	0,0

Five SASs: $\{(U, L)\}$, $\{(U, C)\}$,

SAS Example



		Bob		
		L	C	R
	U	1,1	1,1	0,0
Ann	M	1,1	0,0	1,0
	D	0,0	0,1	0,0

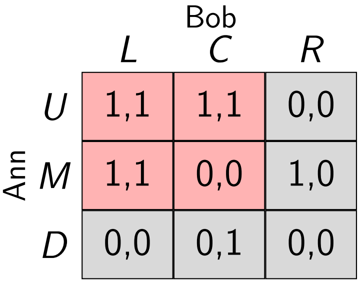
Five SASs: $\{(U, L)\}$, $\{(U, C)\}$, $\{U\} \times \{L, C\}$,

SAS Example

			Bob	
		L	*C*	*R*
Ann	*U*	1,1	1,1	0,0
	M	1,1	0,0	1,0
	D	0,0	0,1	0,0

Five SASs: $\{(U, L)\}$, $\{(U, C)\}$, $\{U\} \times \{L, C\}$, $\{(M, L)\}$, $\{U, M\} \times \{L\}$,

SAS Example



		Bob		
		L	C	R
Ann	U	1,1	1,1	0,0
	M	1,1	0,0	1,0
	D	0,0	0,1	0,0

Five SASs: $\{(U, L)\}$, $\{(U, C)\}$, $\{U\} \times \{L, C\}$, $\{(M, L)\}$, $\{U, M\} \times \{L\}$, but $\{(U, M)\} \times \{L, C\}$ is not an SAS.

SAS Example

A. Brandenburger and A. Friedenberg. *Self-Admissible Sets*. Journal of Economic Theory, 145 (2010), pgs. 785 - 811.

Each type is associated with an LPS: $t_i \mapsto (\mu_0, \mu_1, \ldots, \mu_{n-1})$ (each μ_i is a probability measure on S_{-i} *with disjoint supports*)

Each type is associated with an *LPS*: $t_i \mapsto (\mu_0, \mu_1, \ldots, \mu_{n-1})$ (each μ_i is a probability measure on S_{-i} *with disjoint supports*)

An LPS $(\mu_0, \ldots, \mu_n)$ on X has *full support* if $\cup_{i=1}^{n}\text{supp}(\mu_i) = X$.

Each type is associated with an *LPS*: $t_i \mapsto (\mu_0, \mu_1, \ldots, \mu_{n-1})$ (each μ_i is a probability measure on S_{-i} *with disjoint supports*)

An LPS $(\mu_0, \ldots, \mu_n)$ on X has *full support* if $\cup_{i=1}^{n}\text{supp}(\mu_i) = X$.

(s_i, t_i) is **rational** provided (i) s_i lexicographically maximizes i's expected payoff under the LPS associated with t_i, **and** (ii) the LPS associated with t_i has full support.

Fix an *LPS* $\vec{\mu} = (\mu_0, \ldots, \mu_n)$

- E is certain: $\mu_0(E) = 1$

Fix an *LPS* $\vec{\mu} = (\mu_0, \ldots, \mu_n)$

- E is certain: $\mu_0(E) = 1$
- E is absolutely certain: $\mu_i(E) = 1$ for all $i = 1, \ldots, n$

Fix an *LPS* $\vec{\mu} = (\mu_0, \ldots, \mu_n)$

- E is certain: $\mu_0(E) = 1$
- E is absolutely certain: $\mu_i(E) = 1$ for all $i = 1, \ldots, n$
- E is *assumed*: there exists k such that $\mu_i(E) = 1$ for all $i \leq k$ and $\mu_i(E) = 0$ for all $k < i < n$.

Fix an *LPS* $\vec{\mu} = (\mu_0, \ldots, \mu_n)$

- E is certain: $\mu_0(E) = 1$
- E is absolutely certain: $\mu_i(E) = 1$ for all $i = 1, \ldots, n$
- E is *assumed*: there exists k such that $\mu_i(E) = 1$ for all $i \leq k$ and $\mu_i(E) = 0$ for all $k < i < n$.

The key notion is **rationality and common assumption of rationality** (RCAR).

Digression on Belief Change, I

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

p_1 All Europeans swans are white.

p_2 The bird caught in the trap is a swan.

p_3 The bird caught in the trap comes from Sweden.

p_4 Sweden is part of Europe.

Thus, the agent believes:

q The bird caught in the trap is white.

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

p_1 All Europeans swans are white.
p_2 The bird caught in the trap is a swan.
p_3 The bird caught in the trap comes from Sweden.
p_4 Sweden is part of Europe.

Thus, the agent believes:

q The bird caught in the trap is white.

Now suppose the rational agent—for example, You—learn that the bird caught in the trap is black ($\neg q$).

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

p_1 All Europeans swans are white.

p_2 The bird caught in the trap is a swan.

p_3 The bird caught in the trap comes from Sweden.

p_4 Sweden is part of Europe.

Thus, the agent believes:

q The bird caught in the trap is white.

Question: How should the agent incorporate $\neg q$ into his belief state to obtain a consistent belief state?

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

p_1 All Europeans swans are white.
p_2 The bird caught in the trap is a swan.
p_3 The bird caught in the trap comes from Sweden.
p_4 Sweden is part of Europe.

Thus, the agent believes:

q The bird caught in the trap is white.

Question: How should the agent incorporate $\neg q$ into his belief state to obtain a consistent belief state?

Problem: Logical considerations alone are insufficient to answer this question! Why??

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Question: How should the agent incorporate $\neg q$ into his belief state to obtain a consistent belief state?

Problem: Logical considerations alone are insufficient to answer this question! Why??

There are several logically consistent ways to incorporate $\neg q$!

Digression on Belief Change, II

What extralogical factors serve to determine what beliefs to give up and what beliefs to retain?

Digression on Belief Change, III

Belief revision is a matter of choice, and the choices are to be made in such a way that:

1. The resulting theory squares with the experience;
2. It is simple; and
3. The choices disturb the original theory as little as possible.

Digression on Belief Change, III

Belief revision is a matter of choice, and the choices are to be made in such a way that:

1. The resulting theory squares with the experience;
2. It is simple; and
3. The choices disturb the original theory as little as possible.

Research has relied on the following related guiding ideas:

1. When accepting a new piece of information, an agent should aim at a minimal change of his old beliefs.
2. If there are different ways to effect a belief change, the agent should give up those beliefs which are least entrenched.

Digression: Belief Revision

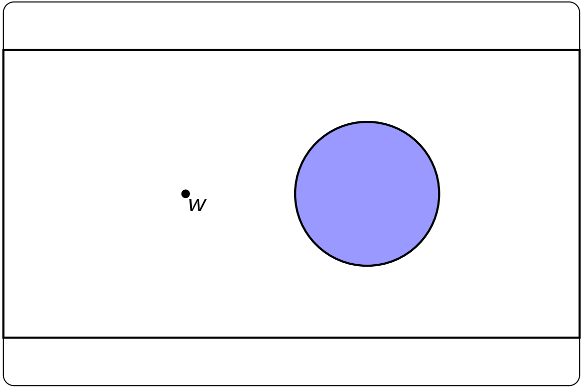
A.P. Pedersen and H. Arló-Costa. *"Belief Revision."*. In Continuum Companion to Philosophical Logic. Continuum Press, 2011.

Hans Rott. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*. Oxford University Press, 2001.

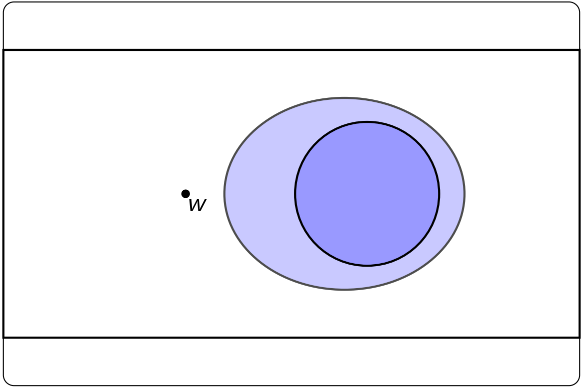
AGM

Let B the set of states representing the prior belief state, and B' the set of states representing the posterior belief state. If E is any subset of B', then $B(E)$ is the set that represents the beliefs state induced by E:

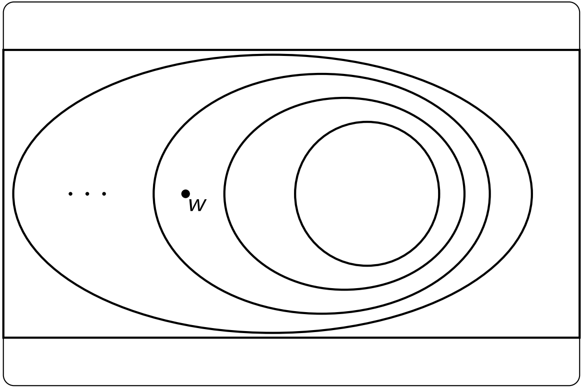
1. For any E, $B(E) \subseteq E$
2. If $E \neq \emptyset$, then $B(E) \neq \emptyset$
3. If $E \cap B \neq \emptyset$, then $B(E) = B \cap E$
4. If $B(E) \cap E \neq \emptyset$, then $B(E \cap F) = B(E) \cap F$



- The agent's (hard) information (i.e., the states consistent with what the agent knows)
- The agent's beliefs (soft information—-the states consistent with what the agent believes)



- The agent's beliefs (soft information—-the states consistent with what the agent believes)
- The agent's "contingency plan": when the stronger beliefs fail, go with the weaker ones.



- The agent's beliefs (soft information—-the states consistent with what the agent believes)
- The agent's "contingency plan": when the stronger beliefs fail, go with the weaker ones.

$\langle W, \{\succeq_i, P_i\}_{i \in N}, \mathbf{s} \rangle$

$\succeq_i$ is a reflexive, transitive and locally connected ordering on W.

$\langle W, \{\succeq_i, P_i\}_{i \in N}, \mathbf{s} \rangle$

$\succeq_i$ is a reflexive, transitive and locally connected ordering on W.

$w \; R_i \; v$ iff $v \in \max_{\succeq_i}(W)$.
$w \approx_i v$ iff $w \succeq_i v$ or $v \succeq_i w$.

$\langle W, \{\succeq_i, P_i\}_{i \in N}, \mathbf{s} \rangle$

$\succeq_i$ is a reflexive, transitive and locally connected ordering on W.

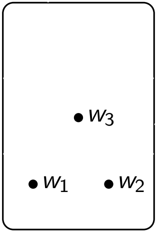
$w \; R_i \; v$ iff $v \in \max_{\succeq_i}(W)$.
$w \approx_i v$ iff $w \succeq_i v$ or $v \succeq_i w$.

$B_{i,x}(E) = \{w \in E \mid \text{ for all } y \in E \cap \{z \mid z \approx_i x\}, w \succeq_i y\}$

$y \succeq_i x$ provided $y \in B(\{x, y\})$

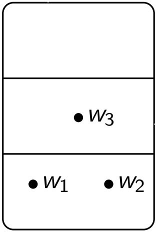
Belief Revision via Plausibility

- $W = \{w_1, w_2, w_3\}$



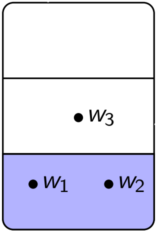
Belief Revision via Plausibility

- $W = \{w_1, w_2, w_3\}$
- $w_1 \preceq w_2$ and $w_2 \preceq w_1$ (w_1 and w_2 are equi-plausbile)
- $w_1 \prec w_3$ ($w_1 \preceq w_3$ and $w_3 \not\preceq w_1$)
- $w_2 \prec w_3$ ($w_2 \preceq w_3$ and $w_3 \not\preceq w_2$)

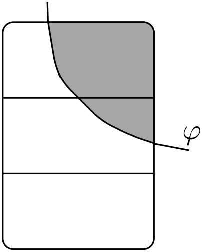


Belief Revision via Plausibility

- $W = \{w_1, w_2, w_3\}$
- $w_1 \preceq w_2$ and $w_2 \preceq w_1$ (w_1 and w_2 are equi-plausbile)
- $w_1 \prec w_3$ ($w_1 \preceq w_3$ and $w_3 \not\preceq w_1$)
- $w_2 \prec w_3$ ($w_2 \preceq w_3$ and $w_3 \not\preceq w_2$)
- $\{w_1, w_2\} \subseteq Min_{\preceq}([w_i])$

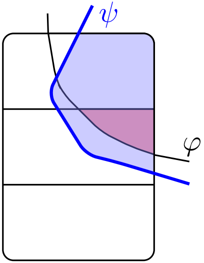


Belief Revision via Plausibility



Conditional Belief: $B^{\varphi}\psi$

Belief Revision via Plausibility



Conditional Belief: $B^{\varphi}\psi$

$$Max_{\preceq}(\llbracket\varphi\rrbracket_{\mathcal{M}}) \subseteq \llbracket\psi\rrbracket_{\mathcal{M}}$$

$E_i^1 = \{x \in W \mid$ for some y such that $y \approx_i x$, not $x \succeq_i y\}$
$= \{x \in W \mid$ not $x \; R_i \; x\}$.

$E_i^{k+1} = \{x \in E_i^k \mid$ for some $y \in E_i^k$ such that $y \approx_i x$, not $x \succeq_i y\}$

E_i^1 is the proposition that player i has at least some false belief

"Even though each of two propositions has maximum degree of belief, one may be believed *more robustly* than the other in the sense that the agent is more disposed to continue believing it in response to new information." (pg. 147, Stalnaker)

$$P_{i,x}(E \mid F) = \frac{P_i(E \cap B_{i,x}(F))}{P_i(B_{i,x}(F))}$$

If $P_{i,x}(F) > 0$, then this coincides with conditional probability. In particular, $P_{i,x}(E) = P_{i,x}(E \mid W)$.

$$P_{i,x}(E \mid F) = \frac{P_i(E \cap B_{i,x}(F))}{P_i(B_{i,x}(F))}$$

If $P_{i,x}(F) > 0$, then this coincides with conditional probability. In particular, $P_{i,x}(E) = P_{i,x}(E \mid W)$.

Definition. An action is **perfectly rational** if it not only maximizes expected utility, but also satisfies a tie-breaking procedure that requires that certain *conditional* expected utilities be maximized as well. The idea is that in cases where two ore more actions maximize expected utility, the agent should consider, in choosing between them, how he should act if he learned he was in error about something. (Stalnaker, pg. 148)

$$EU_{i,x}(s_i \mid E) = \sum_{s_{-i} \in S_{-i}} P_{i,x}([s_{-i}] \mid E) \times u_i((s_i, s_{-i})$$

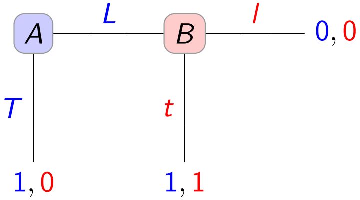
Let $EU_{i,x}(s_i) = EU_{i,x}(s_i \mid W)$

$$EU_{i,x}(s_i \mid E) = \sum_{s_{-i} \in S_{-i}} P_{i,x}([s_{-i}] \mid E) \times u_i((s_i, s_{-i}))$$

Let $EU_{i,x}(s_i) = EU_{i,x}(s_i \mid W)$

$\mathsf{Rat}^0_{i,x} = \mathsf{Rat}_{i,x} = \{s_i \in S_i \mid EU_{i,x}(s_i) \geq EU_{i,x}(s'_i) \text{ for all } s'_i \in S_i\}$
$\mathsf{Rat}^{k+1}_{i,x} = \mathsf{Rat}_{i,x} = \{s_i \in \mathsf{Rat}^k_{i,x} \mid EU_{i,x}(s_i \mid E_i^{k+1}) \geq$
$EU_{i,x}(s'_i \mid E_i^{k+1}) \text{ for all } s'_i \in \mathsf{Rat}^k_{i,x}\}$

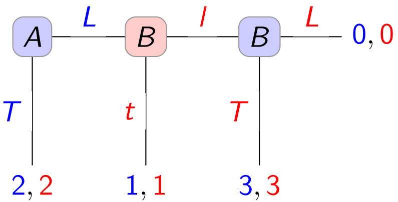
		Bob	
		t	*l*
Ann	*T*	1,0	1,0
	L	1,1	0,0



- Both strategies of both players is rationalizable.
- Only T is perfectly rational for Ann and t is perfectly rational for Bob.

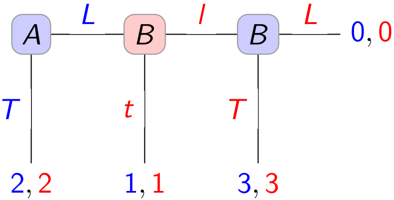
Bob

Ann	*t*	*l*
T	2,2	2,2
LT	1,1	3,3
LL	1,1	0,0

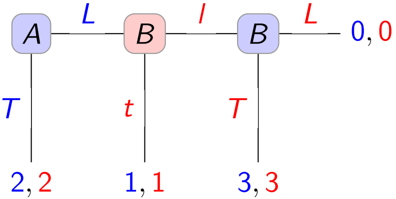


- Suppose that Bob believes that Ann will choose *T* with probability 1; what should he do? This depends on what he thinks Ann would on the hypothesis that his believe about her is mistaken.
- Suppose that if Bob were surprised by her, then he concludes she is irrational, selecting *L* on her second move. Bob's choice of *t* is perfectly rational.

		Bob	
		t	*l*
Ann	*T*	2,2	2,2
	LT	1,1	3,3
	LL	1,1	0,0



- Suppose Ann is sure that Bob will choose t, which is the only perfectly rational choice for Bob. Then, Ann's only rational choice is T.
- So, it might be that Ann and Bob both know each other's beliefs about each other, and are both perfectly rational, but they still fail to coordinate on the optimal outcome for both.



- Perhaps if Bob believed that Ann would choose *L* are her second move then he wouldn't believe she was fully rational, *but it is not suggested that he believes this*.
- Divide Ann's strategy *T* into two *TT*: *T*, and I *would* choose *T* again on the second move if I were faced with that choice" and *TL*: "*T*, but I *would* choose *L* on the second move..."
- Of these two only *TT* is rational
- But if Bob learned he was wrong, he would conclude she chooses *LL*.

“To think there is something incoherent about this combination of beliefs and belief revision policy is to confuse epistemic with causal counterfactuals—it would be like thinking that because I believe that if Shakespeare hadn’t written Hamlet, it would have never been written by anyone, I must therefore be disposed to conclude that Hamlet was never written, were I to learn that Shakespeare was in fact not its author”

(pg. 152, Stalnaker)