Epistemic Game Theory

Lecture 11

Eric Pacuit

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

April 24, 2014

Reasoning Based Expected Utility Procedure

R. Cubitt and R. Sugden. *The reasoning-based expected utility procedure*. Games and Economic Behavior, 2010.

Reasoning-Based Solution Concepts

Reasoning-Based Solution Concepts

A **categorization** is a *ternary* partition of the players choices (rather than a binary partition of what is in and what is out):

Reasoning-Based Solution Concepts

A **categorization** is a *ternary* partition of the players choices (rather than a binary partition of what is in and what is out): strategies are accumulated, deleted or neither.

Reasoning-Based Solution Concepts

A **categorization** is a *ternary* partition of the players choices (rather than a binary partition of what is in and what is out): strategies are accumulated, deleted or neither.

Example: RBEU (reasoning based expected utility):

Reasoning-Based Solution Concepts

A **categorization** is a *ternary* partition of the players choices (rather than a binary partition of what is in and what is out): strategies are accumulated, deleted or neither.

Example: RBEU (reasoning based expected utility):

- accumulate strategies that maximize expected utility for every possibly probability distribution
- delete strategies that do not maximize probability against any probability distribution

Reasoning-Based Solution Concepts

A **categorization** is a *ternary* partition of the players choices (rather than a binary partition of what is in and what is out): strategies are accumulated, deleted or neither.

Example: RBEU (reasoning based expected utility):

- accumulate strategies that maximize expected utility for every possibly probability distribution
- delete strategies that do not maximize probability against any probability distribution
- accumulated strategies must receive positive probability, deleted strategies must receive zero probability

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L, R\}$$
$$S^- = \{B, M_1\}$$

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

$$S^+ = \{L, R\}$$
$$S^- = \{B, M_1\}$$

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

$$S^+ = \{L, R\}$$
$$S^- = \{B, M_1\}$$

RBEU: Example

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

	L	R
U	1,1	1,1
M_1	0,0	1,0
M_2	2,0	0,0
B	0,2	0,0

$$S^+ = \{L\}$$
$$S^- = \{B\}$$

$$S^+ = \{L, R\}$$
$$S^- = \{B, M_1\}$$

$$S^+ = \{L, R\}$$
$$S^- = \{B, M_1\}$$

RBEU Example 2

	l	r
u	1,1	0,0
d	0,0	0,0

RBEU Example 2

	l	r
u	1,1	0,0
d	0,0	0,0

$$S^+ = \{u, l\}$$
$$S^- = \emptyset$$

RBEU Example 2

	l	r
u	1,1	0,0
d	0,0	0,0

	l	r
u	1,1	0,0
d	0,0	0,0

$$S^{+} = \{u, l\}$$
$$S^{-} = \emptyset$$

$$S^{+} = \{u, l\}$$
$$S^{-} = \{d, r\}$$

RBEU Example 3

	L	R
u	1,1	1,0
d	1,0	0,1

RBEU Example 3

	L	R
u	1,1	1,0
d	1,0	0,1

$$S^+ = \{u\}$$
$$S^- = \emptyset$$

RBEU Example 3

	L	R
u	1,1	1,0
d	1,0	0,1

	L	R
u	1,1	1,0
d	1,0	0,1

$$S^+ = \{u\}$$
$$S^- = \emptyset$$

RBEU Example 3

	L	R
u	1,1	1,0
d	1,0	0,1

$$S^+ = \{u\}$$
$$S^- = \emptyset$$

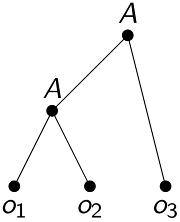
	L	R
u	1,1	1,0
d	1,0	0,1

$$S^+ = \{u\}$$
$$S^- = \emptyset$$

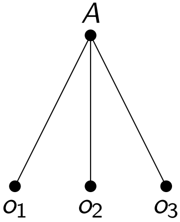
When are two games the *same*?

- Whose point-of-view? (players, modelers)
- Game-theoretic analysis should not depend on "irrelevant" mathematical details
- Different perspectives: transformations, structural, agent

The same decision problem



D_1



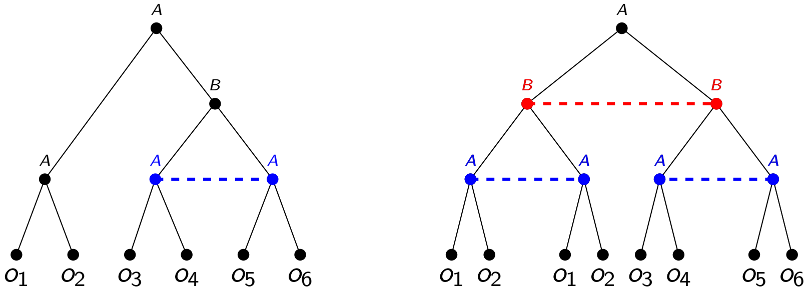
D_2

Thompson Transformations

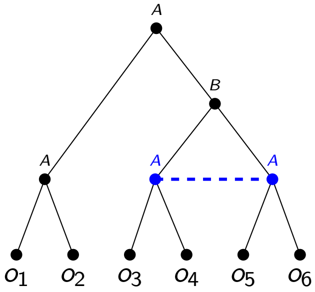
Game-theoretic analysis should not depend on "irrelevant" features of the (mathematical) description of the game.

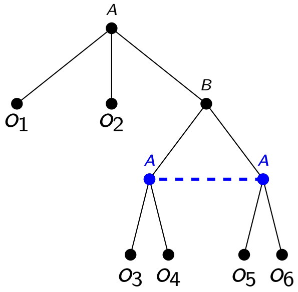
F. B. Thompson. *Equivalence of Games in Extensive Form*. Classics in Game Theory, pgs 36 - 45, 1952.

(Osborne and Rubinstein, pgs. 203 - 212)

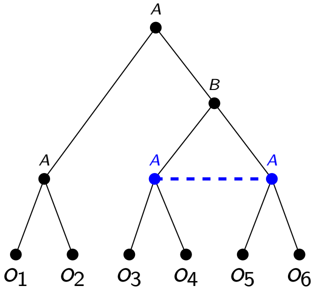


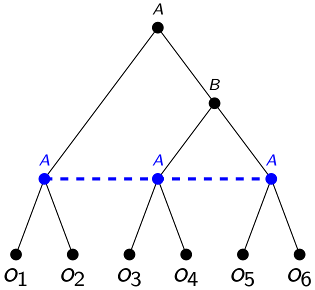
Addition of Superfluous Move



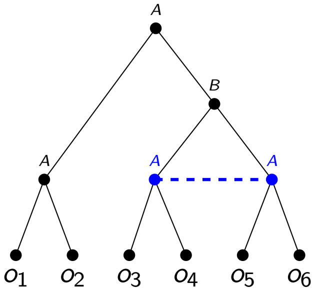


Coalescing of moves





Inflation/deflation



Interchange of moves

Theorem (Thompson) Each of the previous transformations preserves the reduced strategic form of the game. In finite extensive games (without uncertainty between subhistories), if any two games have the same reduced normal form then one can be obtained from the other by a sequence of the four transformations.

Other transformations/game forms

Kholberg and Mertens. *On Strategic Stability of Equilibria*. Econometrica (1986).

Elmes and Reny. *On The Strategic Equivalence of Extensive Form Games*. Journal of Economic Theory (1994).

G. Bonanno. *Set-Theoretic Equivalence of Extensive-Form Games*. IJGT (1992).

Games as Processes

- When are two processes the same?

Games as Processes

- When are two processes the same?
- Extensive games are natural process models which support many familiar modal logics such as *bisimulation*.

Games as Processes

- When are two processes the same?
- Extensive games are natural process models which support many familiar modal logics such as *bisimulation*.
- From this point-of-view, "When are two games the same?" goes tandem with asking "what are appropriate languages for games"

Games as Processes

- When are two processes the same?
- Extensive games are natural process models which support many familiar modal logics such as *bisimulation*.
- From this point-of-view, "When are two games the same?" goes tandem with asking "what are appropriate languages for games"

J. van Benthem. *Extensive Games as Process Models*. IJGT, 2001.

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.

- Strategies as programs: "if she plays this, then I will play that" $[((\mathbf{turn_E})?; \sigma) \cup (\mathbf{turn_A}?; \tau))^*](\mathbf{end} \to p)$

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.

- Strategies as programs: "if she plays this, then I will play that"
$[((\mathbf{turn_E})?;\sigma) \cup (\mathbf{turn_A}?;\tau))^*](\mathbf{end} \rightarrow p)$

- $\Box_A\varphi$: "A has a strategy to ensure that φ is true"

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.

- Strategies as programs: "if she plays this, then I will play that" $[((\mathbf{turn_E})?;\sigma) \cup (\mathbf{turn_A}?;\tau))^*](\mathbf{end} \rightarrow p)$

- $\Box_A\varphi$: "A has a strategy to ensure that φ is true"

- Knoweldge/beliefs:

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.
- Strategies as programs: "if she plays this, then I will play that" $[((\mathbf{turn_E})?;\sigma) \cup (\mathbf{turn_A}?;\tau))^*](\mathbf{end} \rightarrow p)$
- $\Box_A\varphi$: "A has a strategy to ensure that φ is true"
- Knoweldge/beliefs:
$K_E(\langle a\rangle p \vee \langle b\rangle p)$, $\neg K_E\langle a\rangle p \wedge \neg K_E\langle b\rangle p$

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.

- Strategies as programs: "if she plays this, then I will play that"
$[((\mathbf{turn_E})?;\sigma) \cup (\mathbf{turn_A}?;\tau))^*](\mathbf{end} \rightarrow p)$

- $\Box_A\varphi$: "A has a strategy to ensure that φ is true"

- Knoweldge/beliefs:
$K_E(\langle a\rangle p \vee \langle b\rangle p)$, $\neg K_E\langle a\rangle p \wedge \neg K_E\langle b\rangle p$
$K_A\Box_E\varphi$ vs. $\Box_E K_A\varphi$

- $[a \cup b]\langle c \cup d\rangle p$: "for each choice between a or b there is a choice between c or d ending in a p-state.

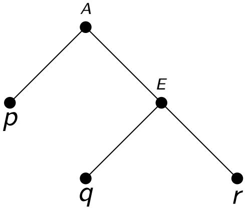
- Strategies as programs: "if she plays this, then I will play that"
  $[((\mathbf{turn_E})?;\sigma) \cup (\mathbf{turn_A}?;\tau))^*](\mathbf{end} \rightarrow p)$

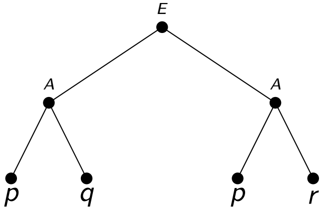
- $\Box_A\varphi$: "A has a strategy to ensure that φ is true"

- Knoweldge/beliefs:
  $K_E(\langle a\rangle p \vee \langle b\rangle p)$, $\neg K_E\langle a\rangle p \wedge \neg K_E\langle b\rangle p$
  $K_A\Box_E\varphi$ vs. $\Box_E K_A\varphi$

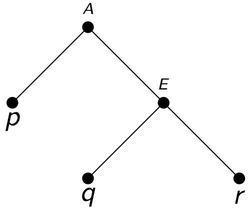
- preferences, ...

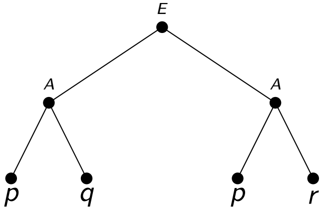
Action and Powers





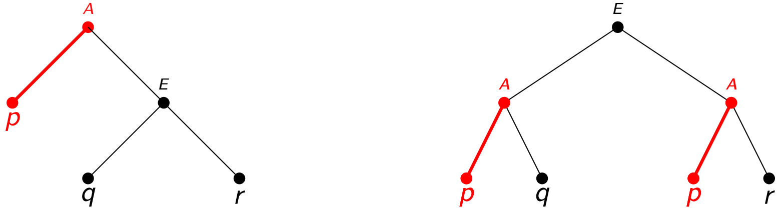
Action and Powers





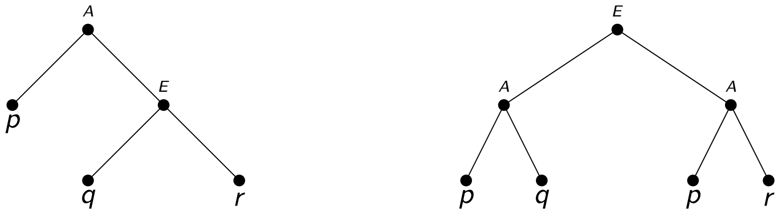
A can force $\{p\}, \{q, r\}$, E can force $\{p, q\}$, $\{p, r\}$

Action and Powers



A can force $\{p\}, \{q, r\}$, E can force $\{p, q\}$, $\{p, r\}$

Action and Powers



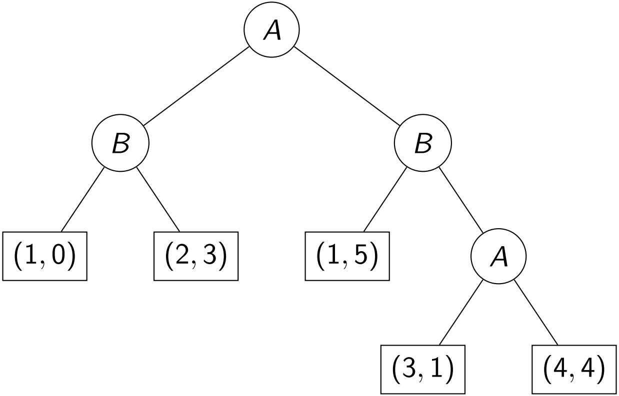
$$(p \wedge (q \vee r)) \leftrightarrow ((p \wedge q) \vee (p \wedge r))$$

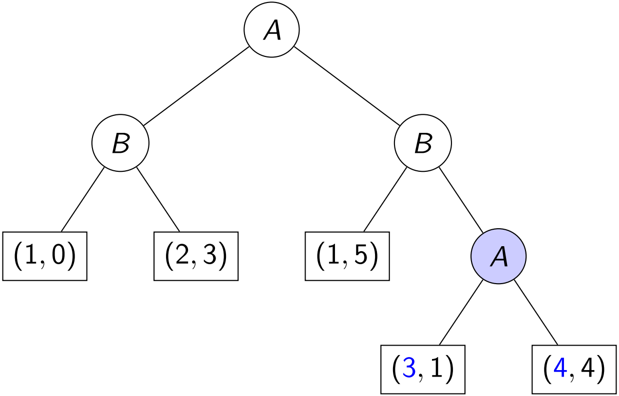
Epistemic Models of Games

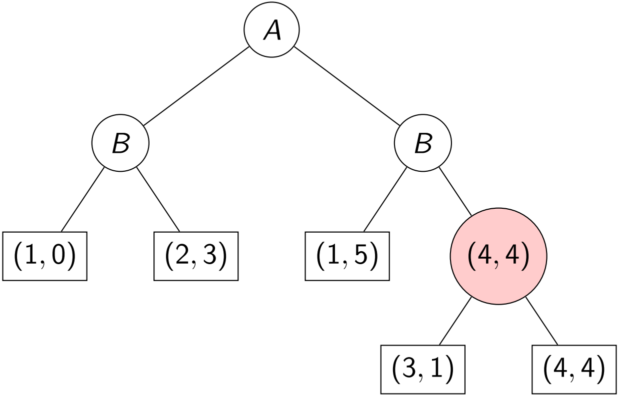
"The aim in giving the general definition of a model is not to propose an original explanatory hypothesis, or any explanatory hypothesis, for the behavior of players in games, but only to provide a descriptive framework for the representation of considerations that are relevant to such explanations, a framework that is as *general* and as *neutral* as we can make it." (pg. 35, Stalnaker)

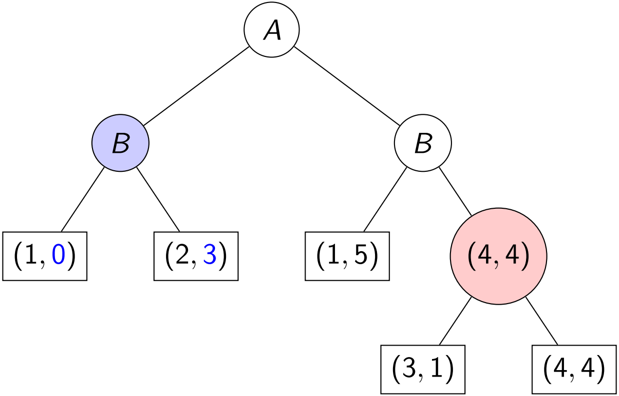
Backwards Induction

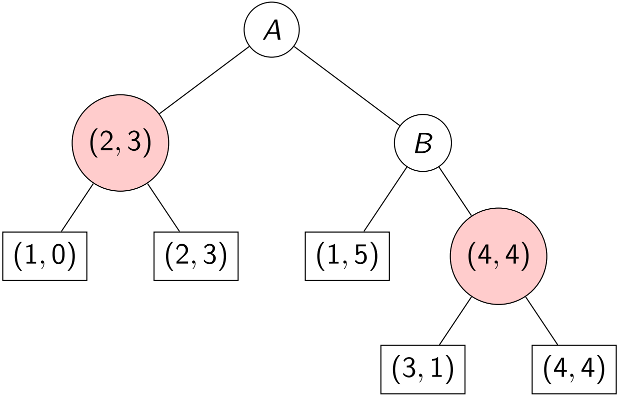
Invented by Zermelo, Backwards Induction is an iterative algorithm for "solving" and extensive game.


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

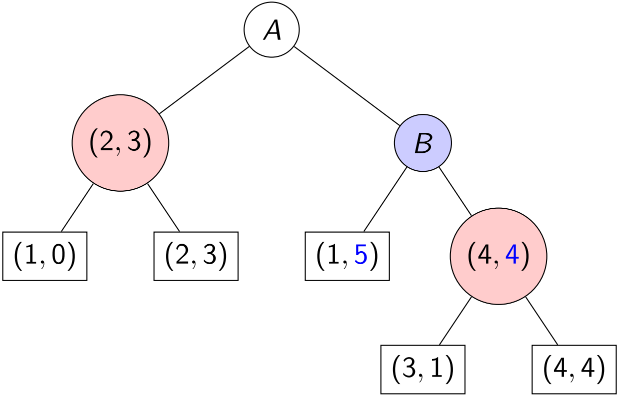

A
B
B
(1, 0)
(2, 3)
(1, 5)
A
(3, 1)
(4, 4)

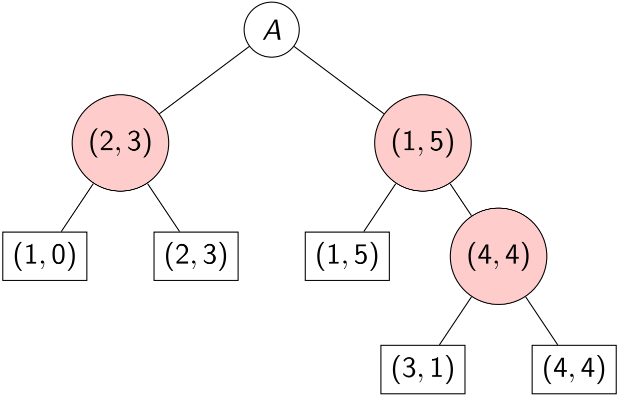

A
B
B
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)


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

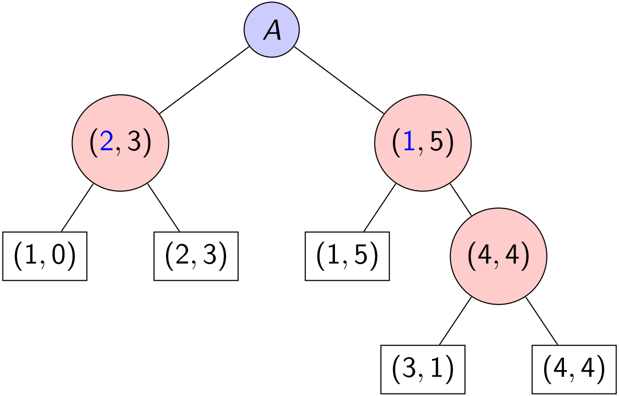


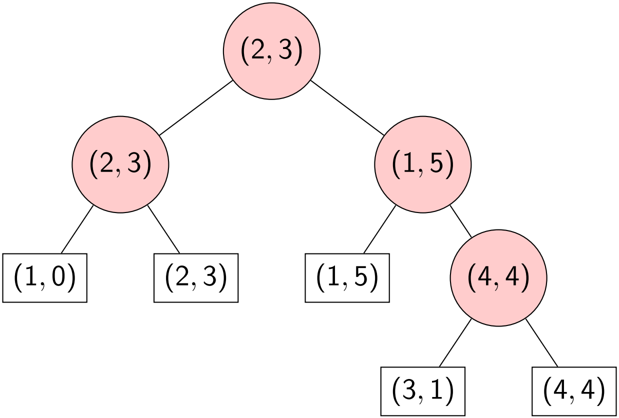
A
(2, 3)
B
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

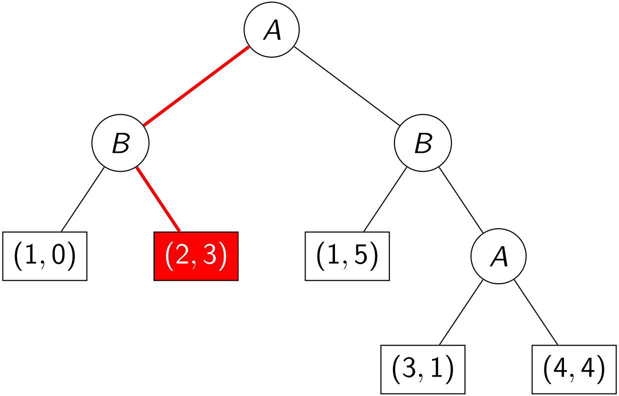

A
(2, 3)
B
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

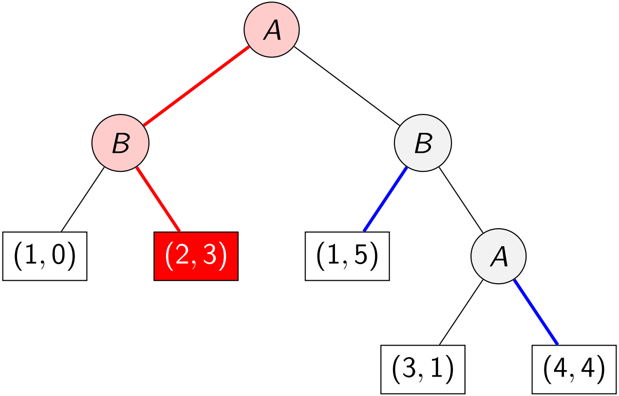


A
(2, 3)
(1, 5)
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)

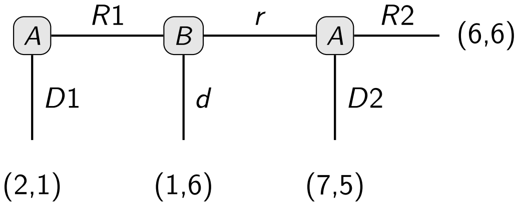

A
(2, 3)
(1, 5)
(1, 0)
(2, 3)
(1, 5)
(4, 4)
(3, 1)
(4, 4)


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

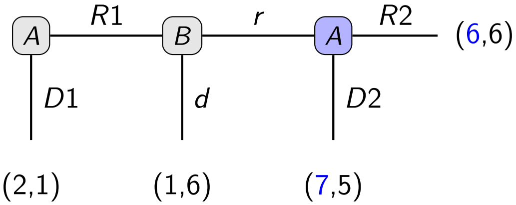

A
B
B
(1, 0)
(2, 3)
(1, 5)
A
(3, 1)
(4, 4)


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

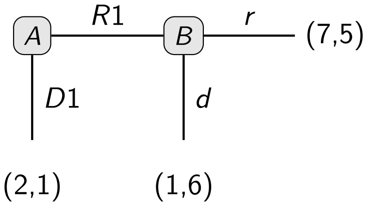
BI Puzzle



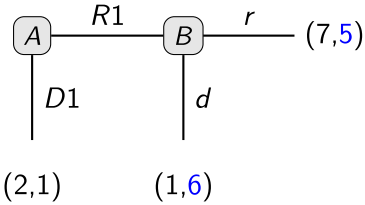
BI Puzzle



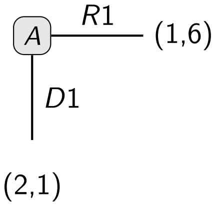
BI Puzzle



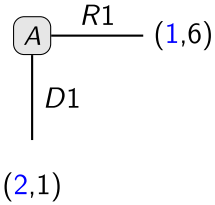
BI Puzzle



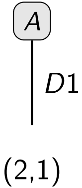
BI Puzzle



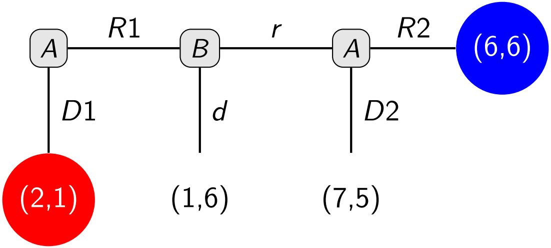
BI Puzzle

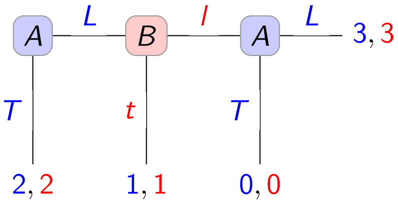


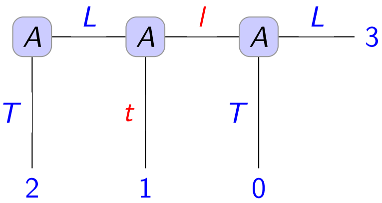
BI Puzzle

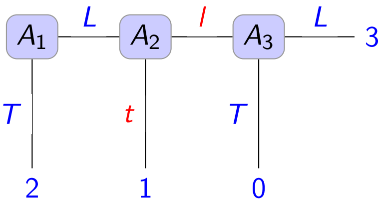


BI Puzzle

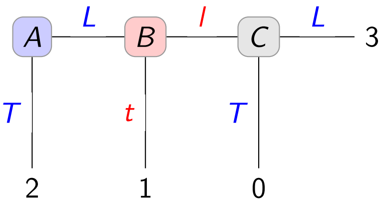


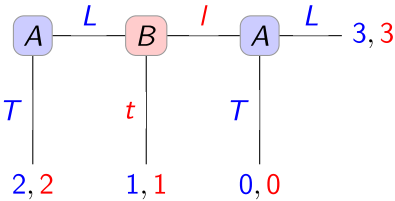

A
L
B
l
A
L
3, 3
T
t
T
2, 2
1, 1
0, 0


A
L
A
l
A
L
3
T
t
T
2
1
0

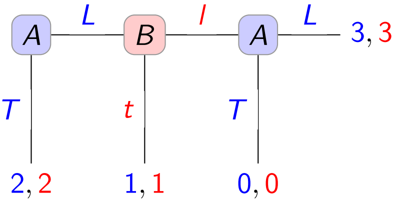


A_1
L
A_2
l
A_3
L
3
T
t
T
2
1
0


A
L
B
l
C
L
3
T
t
T
2
1
0


A
L
B
l
A
L
3, 3
T
t
T
2, 2
1, 1
0, 0

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



R. Aumann. *Backwards induction and common knowledge of rationality*. Games and Economic Behavior, 8, pgs. 6 - 19, 1995.

R. Stalnaker. *Knowledge, belief and counterfactual reasoning in games*. Economics and Philosophy, 12, pgs. 133 - 163, 1996.

J. Halpern. *Substantive Rationality and Backward Induction*. Games and Economic Behavior, 37, pp. 425-435, 1998.

Materially Rational: A player i is materially rational at a state w if every choice actually made is rational.

Substantively Rational: A player i is substantively rational at a state w if the player is materially rational and, in addition, for each *possible* choice, the player *would* have chosen rationally if she had had the opportunity to choose.

Materially Rational: A player i is materially rational at a state w if every choice actually made is rational.

Substantively Rational: A player i is substantively rational at a state w if the player is materially rational and, in addition, for each *possible* choice, the player *would* have chosen rationally if she had had the opportunity to choose.

E.g., Taking keys away from someone who is drunk.

Theorem (Aumann) In any model, if there is common knowledge that the players are substantively rational at state w, the the backward induction solution is played at w.

Two propositions φ and ψ are epistemically independent for player i in world w iff $P_{i,w}(\varphi \mid \psi) = P_{i,w}(\varphi \mid \neg\psi)$ and $P_{i,w}(\psi \mid \varphi) = P_{i,w}(\psi \mid \neg\varphi)$

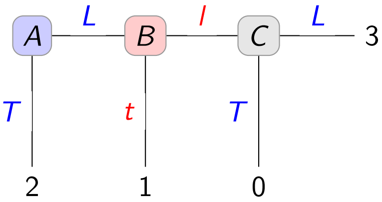
A possible belief revision policy: Information about different players should be epistemically independent.

Theorem (Stalnaker's interpretation of Aumann's theorem) Let G be a game of perfect information in agent form (i.e., players only move once) in which for each player different outcomes have different payoffs. Let $\mathcal{M}$ be a model for G in which it is common belief that all agents are perfectly rational, and that all agents adopt belief revision policies that treat information about different agents as epistemically independent. Then in $\mathcal{M}$, the subgame perfect equilibrium strategy profile is realized.

Theorem (Stalnaker's interpretation of Aumann's theorem) Let G be a game of perfect information in agent form (i.e., players only move once) in which for each player different outcomes have different payoffs. Let $\mathcal{M}$ be a model for G in which it is common belief that all agents are perfectly rational, and that all agents adopt belief revision policies that treat information about different agents as epistemically independent. Then in $\mathcal{M}$, the subgame perfect equilibrium strategy profile is realized.

Theorem (Stalnaker's interpretation of Aumann's theorem) Let G be a game of perfect information in agent form (i.e., players only move once) in which for each player different outcomes have different payoffs. Let $\mathcal{M}$ be a model for G in which it is common belief that all agents are perfectly rational, and that all agents adopt belief revision policies that treat information about different agents as epistemically independent. Then in $\mathcal{M}$, the subgame perfect equilibrium strategy profile is realized.

Theorem (Stalnaker's interpretation of Aumann's theorem) Let G be a game of perfect information in agent form (i.e., players only move once) in which for each player different outcomes have different payoffs. Let $\mathcal{M}$ be a model for G in which it is common belief that all agents are perfectly rational, and that all agents adopt belief revision policies that treat information about different agents as epistemically independent. Then in $\mathcal{M}$, the subgame perfect equilibrium strategy profile is realized.


A
L
B
l
C
L
3
T
t
T
2
1
0

1. If Shakespeare had not written Hamlet, it would never have been written.
2. If Shakespeare didn't write Hamlet, someone else did.

1. is a causal counterfactual, and 2. is an expression of a belief revision policy.

1. General Smith is a shrewd judge of character—he knows (better than I) who is brave and who is not.
2. The general sends only brave men into battle.
3. Private Jones is cowardly.

I believe that (1) Jones would run away if he were sent into battle and (2) if Jones *is* sent into battle, then he won't run away.

1. Ann cheats — she has seen her opponent's cards.
2. Ann has a losing hand, since I have seen both her hand and her opponent's.
3. Ann is rational.

So, I conclude that she will not bet. But how should I revise my beliefs if I learn that Ann did bet?

1. Ann cheats — she has seen her opponent's cards.
2. Ann has a losing hand, since I have seen both her hand and her opponent's.
3. Ann is rational.

So, I conclude that she will not bet. But how should I revise my beliefs if I learn that Ann did bet?

It may be perfectly reasonable for me to be disposed to give up 2.

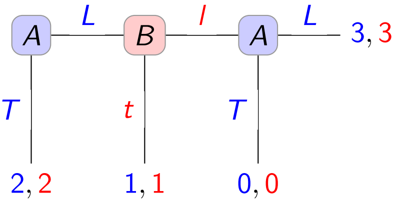
1. Ann cheats — she has seen her opponent's cards.
2. Ann has a losing hand, since I have seen both her hand and her opponent's.
3. Ann is rational.

So, I conclude that she will not bet. But how should I revise my beliefs if I learn that Ann did bet?

It may be perfectly reasonable for me to be disposed to give up 2.

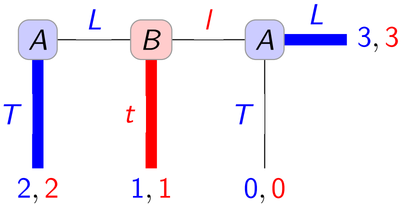
I believe that (1) I Ann *were* to bet, she would lose (since she has a losing hand) and (2) If I were to learn that she *did* bet, I would conclude she will win.

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



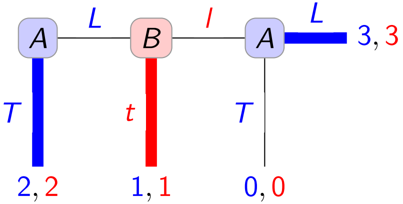
- The backward induction solution is (LL, l)
- Consider a model with a single possible world assigned the profile (TL, t).

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



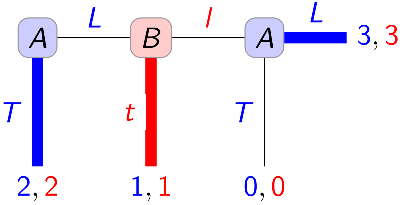
- T is a best response to t, so Ann is materially rational. She is also substantively rational. (Why?)
- Bob doesn't move, so Bob is materially rational. Is he substantively rational?

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



- Is Bob substantively rational? Would t be rational, *if* he had a chance to act?
- Suppose that Bob is disposed to revise his beliefs in such a way that if Ann acted irrationally once, she will act irrationally later in the game.

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



- Bob's belief in a causal counterfactual: Ann would choose L on her second move *if* she had a chance to move.
- But we need to ask what would Bob believe about Ann *if* he learned that he was wrong about her first choice. This is a question about Bob's belief revision policy.

Models of Extensive Games

Let Γ be a *non-degenerate* extensive game with perfect information. Let Γ_i be the set of nodes controlled by player i.

Models of Extensive Games

Let Γ be a *non-degenerate* extensive game with perfect information. Let Γ_i be the set of nodes controlled by player i.

A strategy profile σ describes the choice for each player i at all vertices where i can choose.

Models of Extensive Games

Let Γ be a *non-degenerate* extensive game with perfect information. Let Γ_i be the set of nodes controlled by player i.

A strategy profile σ describes the choice for each player i at all vertices where i can choose.

Given a vertex v in Γ and strategy profile σ, σ specifies a unique path from v to an end-node.

Models of Extensive Games

Let Γ be a *non-degenerate* extensive game with perfect information. Let Γ_i be the set of nodes controlled by player i.

A strategy profile σ describes the choice for each player i at all vertices where i can choose.

Given a vertex v in Γ and strategy profile σ, σ specifies a unique path from v to an end-node.

$\mathcal{M}(\Gamma) = \langle W, \sim_i, \sigma \rangle$ where $\sigma : W \rightarrow \mathit{Strat}(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation.

Models of Extensive Games

Let Γ be a *non-degenerate* extensive game with perfect information. Let Γ_i be the set of nodes controlled by player i.

A strategy profile σ describes the choice for each player i at all vertices where i can choose.

Given a vertex v in Γ and strategy profile σ, σ specifies a unique path from v to an end-node.

$\mathcal{M}(\Gamma) = \langle W, \sim_i, \sigma \rangle$ where $\sigma : W \to \mathit{Strat}(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation.

If $\sigma(w) = \sigma$, then $\sigma_i(w) = \sigma_i$ and $\sigma_{-i}(w) = \sigma_{-i}$

Models of Extensive Games

Let Γ be a *non-degenerate* extensive game with perfect information. Let Γ_i be the set of nodes controlled by player i.

A strategy profile σ describes the choice for each player i at all vertices where i can choose.

Given a vertex v in Γ and strategy profile σ, σ specifies a unique path from v to an end-node.

$\mathcal{M}(\Gamma) = \langle W, \sim_i, \sigma \rangle$ where $\sigma : W \to Strat(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation.

If $\sigma(w) = \sigma$, then $\sigma_i(w) = \sigma_i$ and $\sigma_{-i}(w) = \sigma_{-i}$

(A1) If $w \sim_i w'$ then $\sigma_i(w) = \sigma_i(w')$.

Rationality

$h_i^v(\sigma)$ denote "i's payoff if σ is followed from node v"

Rationality

$h_i^v(\sigma)$ denote "i's payoff if σ is followed from node v"

i **is rational at** v **in** w provided for all strategies $s_i \neq \sigma_i(w)$, $h_i^v(\sigma(w')) \geq h_i^v((\sigma_{-i}(w'), s_i))$ for some $w' \in [w]_i$.

Substantive Rationality

i is **substantively rational** in state w if i is rational at a vertex v in w of every vertex in $v \in \Gamma_i$

Stalnaker Rationality

For every vertex $v \in \Gamma_i$, *if i were to actually reach v, then what he would do in that case would be rational*.

Stalnaker Rationality

For every vertex $v \in \Gamma_i$, *if i were to actually reach v, then what he would do in that case would be rational*.

$f : W \times \Gamma_i \to W$, $f(w, v) = w'$, then w' is the "closest state to w where the vertex v is reached.

Stalnaker Rationality

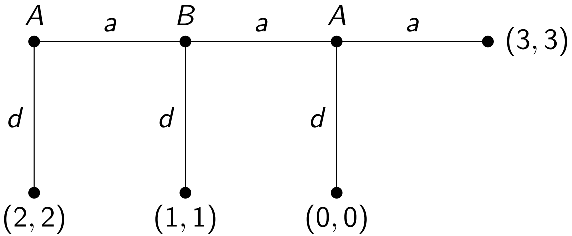
For every vertex $v \in \Gamma_i$, *if i were to actually reach v, then what he would do in that case would be rational*.

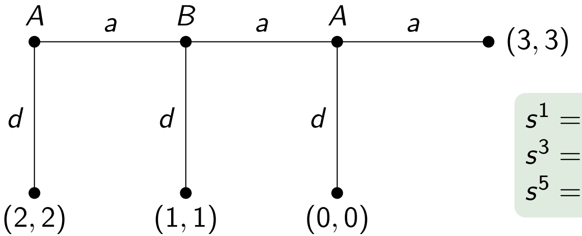
$f : W \times \Gamma_i \to W$, $f(w, v) = w'$, then w' is the "closest state to w where the vertex v is reached.

(F1) v is reached in $f(w, v)$ (i.e., v is on the path determined by $\sigma(f(w, v))$)

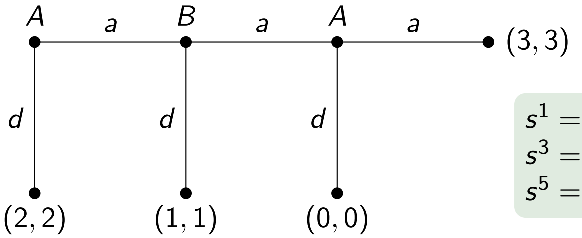
(F2) If v is reached in w, then $f(w, v) = w$

(F3) $\sigma(f(w, v))$ and $\sigma(w)$ agree on the subtree of Γ below v


A
a
B
a
A
a
(3, 3)
d
d
d
(2, 2)
(1, 1)
(0, 0)

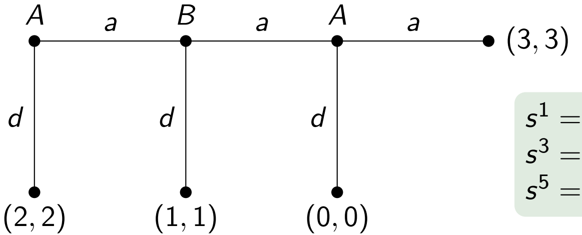


$s^1 = (da, d)$, $s^2 = (aa, d)$,
$s^3 = (ad, d)$, $s^4 = (aa, a)$,
$s^5 = (ad, a)$

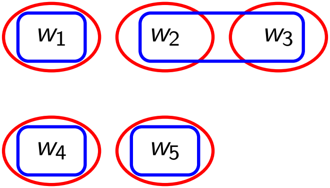


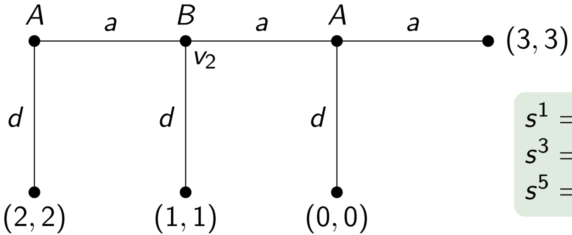
$s^1 = (da, d)$, $s^2 = (aa, d)$,
$s^3 = (ad, d)$, $s^4 = (aa, a)$,
$s^5 = (ad, a)$

- $W = \{w_1, w_2, w_3, w_4, w_5\}$ with $\sigma(w_i) = s^i$
- $[w_i]_A = \{w_i\}$ for $i = 1, 2, 3, 4, 5$
- $[w_i]_B = \{w_i\}$ for $i = 1, 4, 5$ and $[w_2]_B = [w_3]_B = \{w_2, w_3\}$

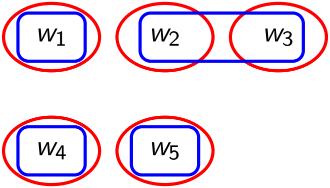


$s^1 = (da, d),\ s^2 = (aa, d),$
$s^3 = (ad, d),\ s^4 = (aa, a),$
$s^5 = (ad, a)$

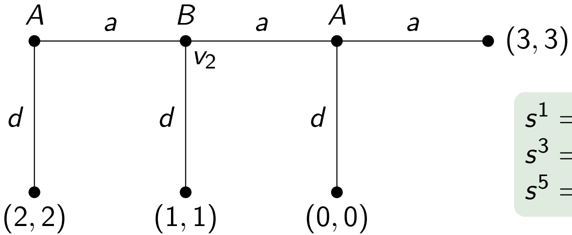




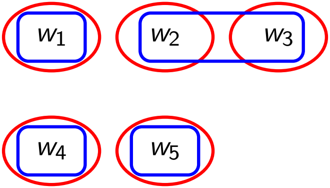
$s^1 = (da, d)$, $s^2 = (aa, d)$,
$s^3 = (ad, d)$, $s^4 = (aa, a)$,
$s^5 = (ad, a)$



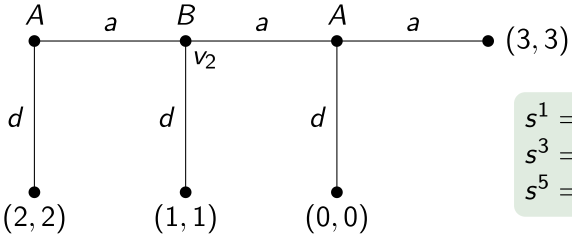
It is **common knowledge** at w_1 that if vertex v_2 were reached, Bob would play down.



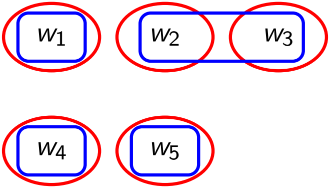
$s^1 = (da, d)$, $s^2 = (aa, d)$,
$s^3 = (ad, d)$, $s^4 = (aa, a)$,
$s^5 = (ad, a)$



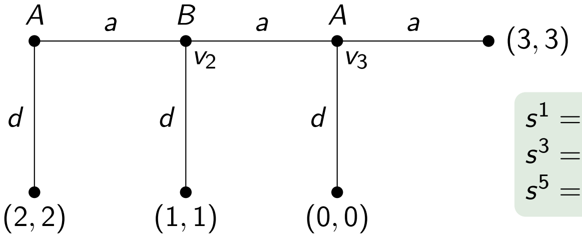
Bob is not rational at v_2 in w_1



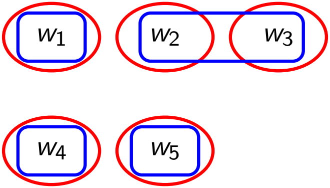
$s^1 = (da, d)$, $s^2 = (aa, d)$,
$s^3 = (ad, d)$, $s^4 = (aa, a)$,
$s^5 = (ad, a)$



Bob is rational at v_2 in w_2



$s^1 = (da, d)$, $s^2 = (aa, d)$,
$s^3 = (ad, d)$, $s^4 = (aa, a)$,
$s^5 = (ad, a)$



Note that $f(w_1, v_2) = w_2$ and $f(w_1, v_3) = w_4$, so there is common knowledge of S-rationality at w_1.

Aumann's Theorem: If Γ is a non-degenerate game of perfect information, then in all models of Γ, we have $C(A - Rat) \subseteq BI$

Stalnaker's Theorem: There exists a non-degenerate game Γ of perfect information and an extended model of Γ in which the selection function satisfies F1-F3 such that $C(S - Rat) \not\subseteq BI$.

Aumann's Theorem: If Γ is a non-degenerate game of perfect information, then in all models of Γ, we have $C(A - Rat) \subseteq BI$

Stalnaker's Theorem: There exists a non-degenerate game Γ of perfect information and an extended model of Γ in which the selection function satisfies F1-F3 such that $C(S - Rat) \not\subseteq BI$.

Revising beliefs during play:

"Although it is common knowledge that Ann would play across if v_3 were reached, if Ann were to play across at v_1, Bob would consider it possible that Ann would play down at v_3"

F4. For all players i and vertices v, if $w' \in [f(w, v)]_i$ then there exists a state $w'' \in [w]_i$ such that $\sigma(w')$ and $\sigma(w'')$ agree on the subtree of Γ below v.

Theorem (Halpern). If Γ is a non-degenerate game of perfect information, then for every extended model of Γ in which the selection function satisfies F1-F4, we have $C(S - Rat) \subseteq BI$. Moreover, there is an extend model of Γ in which the selection function satisfies F1-F4.

J. Halpern. *Substantive Rationality and Backward Induction*. Games and Economic Behavior, 37, pp. 425-435, 1998.