Epistemic Game Theory

Lecture 9

Eric Pacuit

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

April 10, 2014

> “Analysis of strategic economic situations requires us, implicitly or explicitly, to maintain as plausible certain psychological hypotheses. They hypothesis that real economic agents universally recognize the salience of Nash equilibria may well be less accurate than, for example, the hypothesis that agents attempt to “out-smart” or “second-guess” each other, believing that their opponents do likewise.” (pg. 1010)

B. D. Bernheim. *Rationalizable Strategic Behavior*. Econometrica, 52:4, pgs. 1007 - 1028, 1984.

"The rules of a game and its numerical data are seldom sufficient for logical deduction alone to single out a unique choice of strategy for each player. *To do so one requires either richer information (such as institutional detail or perhaps historical precedent for a certain type of behavior) or bolder assumptions about how players choose strategies.* Putting further restrictions on strategic choice is a complex and treacherous task. But one's intuition frequently points to patterns of behavior that cannot be isolated on the grounds of consistency alone."

(pg. 1035)

D. G. Pearce. *Rationalizable Strategic Behavior*. Econometrica, 52, 4, pgs. 1029 - 1050, 1984.

What are the consequences of assuming that the players are *rational* and there is common belief of rationality?

"...a decision-maker has a subjective probability opinion with respect to all of the unknown contingencies affecting his payoffs. In particular in a simultaneous-move two-person game, the player whom we are advising is assumed to have an opinion about the major contingency faced, namely what the opposing player is likely to do. If I think my opponent will choose strategy i ($i = 1, \ldots, n$) with probability p_i, I will choose any strategy j maximizing $\sum_{i=1}^{n} p_i u_{ij}$, where u is the utility to me of the situation in which my opponent has chosen i and I have chosen j."

(pg. 115, Kadane and Larkey)

"It is true that a subjectivist Bayesian will have an opinion not only on his opponent's behavior, but also on his opponent's belief about his own behavior, his opponent's belief about his belief about his opponent's behavior, etc. (He also has opinions about the phase of the moon, tomorrow's weather and the winner of the next Superbowl).

"It is true that a subjectivist Bayesian will have an opinion not only on his opponent's behavior, but also on his opponent's belief about his own behavior, his opponent's belief about his belief about his opponent's behavior, etc. (He also has opinions about the phase of the moon, tomorrow's weather and the winner of the next Superbowl). *However, in a single-play game, all aspects of his opinion except his opinion about his opponent's behavior are irrelevant, and can be ignored in the analysis by integrating them out of the joint opinion.*" (KL, pg. 239, my emphasis)

Theorem. Assume that there is a common prior and that for all w, for all $i \in N$, $\Pi_i(w) \subseteq \{v \mid \mathbf{s}_i(v) = \mathbf{s}_i(w)\}$. If each player is Bayes rational at each state of the world, then the distribution of the action n-tuple $\mathbf{s}$ is a correlated equilibrium.

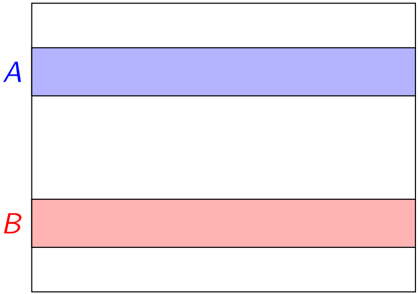
R. Aumann. *Correlated Equilibrium as an Expression of Bayesian Rationality*. Econometrica, 55:1, pgs. 1 - 18, 1987.

Deliberation in Games

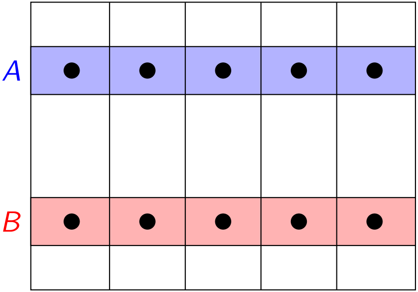
- The Harsanyi-Selten tracing procedure
- Brian Skyrms' models of "dynamic deliberation"
- Ken Binmore's analysis using Turing machines to "calculate" the rational choice
- Robin Cubitt and Robert Sugden's "reasoning based expected utility procedure"
- Johan van Benthem et col.'s "virtual rationality *announcements*"

Different frameworks, common thought: *the "rational solutions" of a game are the result of individual deliberation about the "rational" action to choose.*

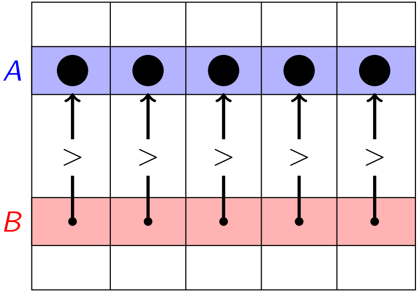
Dominance Reasoning



Dominance Reasoning



Dominance Reasoning

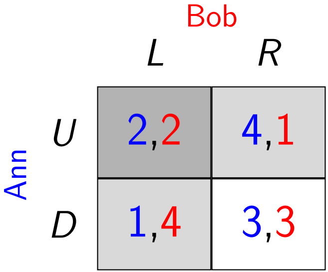


		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

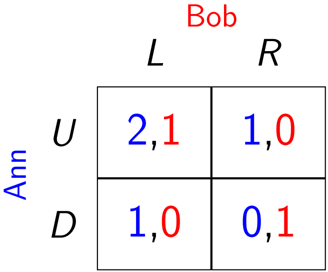
		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2



	L	*R*
U	2,2	4,1
D	1,4	3,3

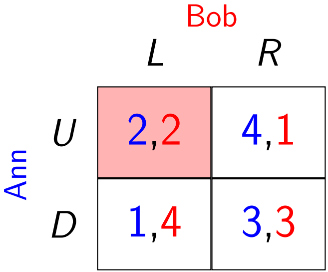
Game 1



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

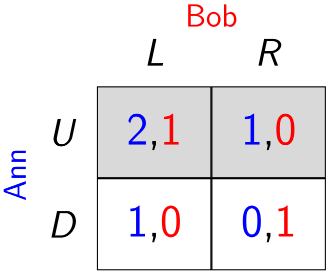
Game 2

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



		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

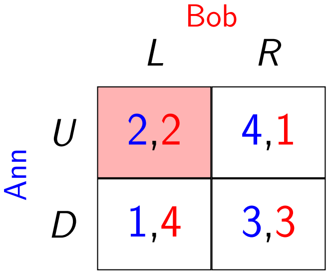


		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2

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

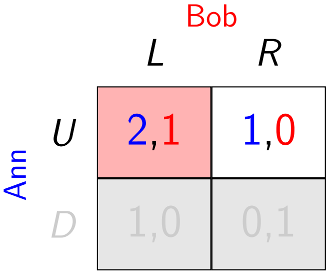
Game 2: U strictly dominates D



Bob

		L	R
Ann	U	2,2	4,1
	D	1,4	3,3

Game 1



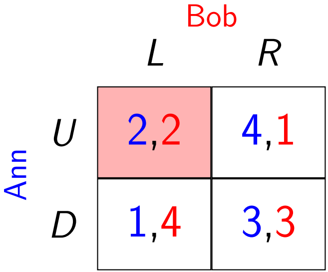
Bob

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

Game 2

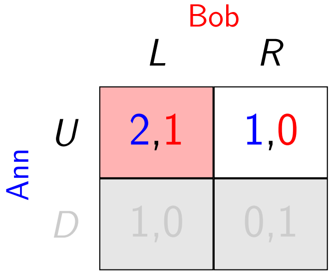
Game 1: U strictly dominates D and L strictly dominates R.

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



		Bob	
		L	R
Ann	U	2,2	4,1
	D	1,4	3,3

Game 1



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

Game 2

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

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

Theorem. In all models where the players are *rational* and there is *common belief of rationality*, the players choose strategies that survive iterative removal of strictly dominated strategies (and, conversely...).

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$
$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$
$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

$s, s' \in S_i$, s **strictly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \quad u_i(s, s_{-i}) > u_i(s', s_{-i})$$

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$
$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

$s, s' \in S_i$, s **strictly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \quad u_i(s, s_{-i}) > u_i(s', s_{-i})$$

$p \in \Delta(X)$, s is a **best response** to p with respect to X provided

$$\forall s' \in S_i, \quad EU(s, p) \geq EU(s', p)$$

		L	Bob	R
	U	5,$*$		1,$*$
Ann	M	1,$*$		5,$*$
	D	2,$*$		2,$*$

D is strictly dominated by $(0.5U, 0.5M)$.

Bob

		L	R
	U	2,*	-1,*
Ann	M	0,*	0,*
	D	-1,*	2,*

M is never a best response: if $p(L) > 1/2$ then U strictly dominates M, if $p(L) < 1/2$, then D strictly dominates M.

Strict Dominance and MEU

Proposition. Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a strategic game and $X \subseteq S_{-i}$. A strategy $s_i \in S_i$ is strictly dominated (possibly by a mixed strategy) with respect to X iff there is no probability measure $p \in \Delta(X)$ such that s_i is a best response to p.

Suppose that $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N}\rangle$ is a finite strategic game.

Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Suppose that $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Let $p \in \Delta(X)$ be any probability measure. Then,

$$\forall s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) \geq p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

$$\exists s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) > p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Let $p \in \Delta(X)$ be any probability measure. Then,

$$\forall s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) \geq p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

$$\exists s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) > p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Hence,

$$\sum_{s_{-i}\in S_{-i}} p(s_{-i}) \cdot u_i(s_i', s_{-i}) > \sum_{s_{-i}\in S_{-i}} p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Let $p \in \Delta(X)$ be any probability measure. Then,

$$\forall s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) \geq p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

$$\exists s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) > p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Hence,

$$\sum_{s_{-i}\in S_{-i}} p(s_{-i}) \cdot u_i(s_i', s_{-i}) > \sum_{s_{-i}\in S_{-i}} p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

So, $EU(s_i', p) > EU(s_i, p)$: s_i is not a best response to p.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two-player game.
(Let $U_i : \Delta(S_1) \times \Delta(S_2) \to \mathbb{R}$ be the expected utility for i)

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two-player game.
(Let $U_i : \Delta(S_1) \times \Delta(S_2) \rightarrow \mathbb{R}$ be the expected utility for i)

Suppose that $\alpha \in \Delta(S_1)$ is not a best response to any $p \in \Delta(S_2)$.

$$\forall p \in \Delta(S_2) \;\; \exists q \in \Delta(S_1), \quad U_1(q, p) > U_1(\alpha, p)$$

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

Let $G = \langle S_1, S_2, u_1, u_2\rangle$ be a two-player game.
(Let $U_i : \Delta(S_1) \times \Delta(S_2) \to \mathbb{R}$ be the expected utility for i)

Suppose that $\alpha \in \Delta(S_1)$ is not a best response to any $p \in \Delta(S_2)$.

$$\forall p \in \Delta(S_2) \;\; \exists q \in \Delta(S_1), \quad U_1(q, p) > U_1(\alpha, p)$$

We can define a function $b : \Delta(S_2) \to \Delta(S_1)$ where, for each $p \in \Delta(S_2)$, $U_1(b(p), p) > U_1(\alpha, p)$.

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

Consider the game $G' = \langle S_1, S_2, \overline{u}_1, \overline{u}_2 \rangle$ where

$$\overline{u}_1(s_1, s_2) = u_1(s_1, s_2) - U_1(\alpha, s_2) \text{ and } \overline{u}_2(s_1, s_2) = -\overline{u}_1(s_1, s_2)$$

Consider the game $G' = \langle S_1, S_2, \overline{u}_1, \overline{u}_2 \rangle$ where

$$\overline{u}_1(s_1, s_2) = u_1(s_1, s_2) - U_1(\alpha, s_2) \text{ and } \overline{u}_2(s_1, s_2) = -\overline{u}_1(s_1, s_2)$$

By the minimax theorem, there is a Nash equilibrium (p_1^*, p_2^*) such that for all $m \in \Delta(S_2)$,

$$\overline{U}_1(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*)$$

Consider the game $G' = \langle S_1, S_2, \overline{u}_1, \overline{u}_2 \rangle$ where

$$\overline{u}_1(s_1, s_2) = u_1(s_1, s_2) - U_1(\alpha, s_2) \text{ and } \overline{u}_2(s_1, s_2) = -\overline{u}_1(s_1, s_2)$$

By the minimax theorem, there is a Nash equilibrium (p_1^*, p_2^*) such that for all $m \in \Delta(S_2)$,

$$\overline{U}_1(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*)$$

We now prove that $\overline{U}_1(b(p_2^*), p_2^*) > 0$:

$$\overline{U}_1(b(p_2^*), p_2^*) \quad = \quad \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) \overline{u}_1(x, y)$$

$$
\begin{aligned}
\overline{U}_1(b(p_2^*), p_2^*) &= \textstyle\sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) \overline{u}_1(x, y) \\
&= \textstyle\sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) [u_1(x, y) - U_1(\alpha, y)]
\end{aligned}
$$

$$
\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\[2ex]
 & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)] \\[2ex]
 & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
 & - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)
\end{array}
$$

$$
\begin{aligned}
\overline{U}_1(b(p_2^*), p_2^*) &= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y)\\[2ex]
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha, y)]\\[2ex]
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y)\\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha, y)\\[2ex]
&= U_1(b(p_2^*), p_2^*)\\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha, y)
\end{aligned}
$$

$$
\begin{aligned}
\overline{U}_1(b(p_2^*), p_2^*) &= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha, y)] \\
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
&= U_1(b(p_2^*), p_2^*) \\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
&> U_1(\alpha, p_2^*) - \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)
\end{aligned}
$$

$$
\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha, y)] \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(b(p_2^*), p_2^*) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& > & U_1(\alpha, p_2^*) - \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(\alpha, p_2^*) - \sum_{x\in S_1} b(p_2^*)(x)\sum_{y\in S_2} p_2^*(y)U_1(\alpha,y)
\end{array}
$$

$$
\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)] \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(b(p_2^*), p_2^*) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& > & U_1(\alpha, p_2^*) - \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(\alpha, p_2^*) - \sum_{x\in S_1} b(p_2^*)(x)\sum_{y\in S_2} p_2^*(y)U_1(\alpha,y) \\
& = & U_1(\alpha, p_2^*) - U_1(\alpha, p_2^*)\cdot\sum_{x\in S_1} b(p_2^*)(x)
\end{array}
$$

$$\begin{aligned}
\overline{U}_1(b(p_2^*), p_2^*) &= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y)\\
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)]\\
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y)\\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)\\
&= U_1(b(p_2^*), p_2^*)\\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)\\
&> U_1(\alpha, p_2^*) - \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)\\
&= U_1(\alpha, p_2^*) - \textstyle\sum_{x\in S_1} b(p_2^*)(x)\sum_{y\in S_2} p_2^*(y)U_1(\alpha,y)\\
&= U_1(\alpha, p_2^*) - U_1(\alpha, p_2^*)\cdot\textstyle\sum_{x\in S_1} b(p_2^*)(x)\\
&= U_1(\alpha, p_2^*) - U_1(\alpha, p_2^*) = 0
\end{aligned}$$

Hence, for all $m \in \Delta(S_2)$ we have

$$\overline{U}(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*) > 0$$

Hence, for all $m \in \Delta(S_2)$ we have

$$\overline{U}(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*) > 0$$

which implies for all $m \in \Delta(S_2)$, $U_1(p_1^*, m) > U_1(\alpha, m)$, and so α is strictly dominated by p_1^*.

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- Note that y is not strictly dominated for Charles.

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- Note that y is not strictly dominated for Charles.
- It is easy to find a probability measure $p \in \Delta(S_A \times S_B)$ such that y is a best response to p. Suppose that $p(u, l) = p(d, r) = \frac{1}{2}$. Then, $EU(x, p) = EU(z, p) = 1.5$ while $EU(y, p) = 2$.

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- Note that y is not strictly dominated for Charles.
- It is easy to find a probability measure $p \in \Delta(S_A \times S_B)$ such that y is a best response to p. Suppose that $p(u, l) = p(d, r) = \frac{1}{2}$. Then, $EU(x, p) = EU(z, p) = 1.5$ while $EU(y, p) = 2$.
- However, there is no probability measure $p \in \Delta(S_A \times S_B)$ such that y is a best response to p and $p(u, l) = p(u) \cdot p(l)$.

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- To see this, suppose that a is the probability assigned to u and b is the probability assigned to l. Then, we have:
  - The expected utility of y is $2ab + 2(1-a)(1-b)$;
  - The expected utility of x is $3ab + 3a(1-b) = 3a(b + (1-b)) = 3a$; and
  - The expected utility of z is $3(1-a)b + 3(1-a)(1-b) = 3(1-a)(b + (1-b)) = 3(1-a)$.

Given a sequence of sets of strategies $S_1, \dots, S_n$ and $s \in S_1 \times \cdots \times S_n$, the following is standard notation:

- $s_{-i} \ := \ (s_1, \dots, s_{i-i}, s_{i+1}, \dots, s_n)$
- $S_{-i} = S_1 \times \cdots \times S_{i-i} \times S_{i+1} \times \cdots S_n$

Given a sequence of sets of strategies $S_1, \ldots, S_n$ and $s \in S_1 \times \cdots \times S_n$, the following is standard notation:

- $s_{-i} \;:=\; (s_1, \ldots, s_{i-i}, s_{i+1}, \ldots, s_n)$
- $S_{-i} = S_1 \times \cdots \times S_{i-i} \times S_{i+1} \times \cdots S_n$

We say that $G = (S_1, \ldots S_n)$ is a **restriction** of a game $H = (T_1, \ldots, T_n, u_1, \ldots, u_n)$ provided $S_i \subseteq T_i$ for all $i = 1, \ldots n$.

Given a sequence of sets of strategies $S_1, \ldots, S_n$ and $s \in S_1 \times \cdots \times S_n$, the following is standard notation:

- $s_{-i} \ := \ (s_1, \ldots, s_{i-i}, s_{i+1}, \ldots, s_n)$
- $S_{-i} = S_1 \times \cdots \times S_{i-i} \times S_{i+1} \times \cdots S_n$

We say that $G = (S_1, \ldots S_n)$ is a **restriction** of a game $H = (T_1, \ldots, T_n, u_1, \ldots, u_n)$ provided $S_i \subseteq T_i$ for all $i = 1, \ldots n$.

A restriction G *where each* S_i *is nonempty* is associated with a unique *subgame*, $\overline{G} = (S_1, \ldots, S_n, u'_1, \ldots u'_n)$ where $u'_i = u_i|_{S_1 \times \cdots S_n}$ (each u'_i is the restriction of u_i to the strategies in S_i).

Given a sequence of sets of strategies $S_1, \ldots, S_n$ and $s \in S_1 \times \cdots \times S_n$, the following is standard notation:

- $s_{-i} \ := \ (s_1, \ldots, s_{i-i}, s_{i+1}, \ldots, s_n)$
- $S_{-i} = S_1 \times \cdots \times S_{i-i} \times S_{i+1} \times \cdots S_n$

We say that $G = (S_1, \ldots S_n)$ is a **restriction** of a game $H = (T_1, \ldots, T_n, u_1, \ldots, u_n)$ provided $S_i \subseteq T_i$ for all $i = 1, \ldots n$.

A restriction G *where each* S_i *is nonempty* is associated with a unique *subgame*, $\overline{G} = (S_1, \ldots, S_n, u'_1, \ldots u'_n)$ where $u'_i = u_i|_{S_1 \times \cdots S_n}$ (each u'_i is the restriction of u_i to the strategies in S_i).

A restriction where some S_i are empty is called an *empty restriction*.

Restrictions of a game H can be ordered by the component-wise subset relation:
$G = (S_1 \ldots, S_n) \subseteq (S'_1, \ldots, S'_n) = G'$ iff $S_i \subseteq S'_i$ for all $i = 1, \ldots n$

Beliefs, or Conjectures

Fix a game $H = (T_1, \ldots, T_n, u_1, \ldots, u_n)$

For each player let $\mathcal{B}_i$ be a set of **beliefs** (for now, this is an unspecified set)

Each u_i is associated with a expected payoff function $U_i : S_i \times \mathcal{B}_i \to \mathbb{R}$.

A belief $\mathcal{B}_i$ of player i in H can be **narrowed** to any restriction G of H. This narrowing of H to G is denoted: $\mathcal{B}_i \dot{\cap} G$

We call the pair $(\mathcal{B}, \dot{\cap})$ a belief structure in the game H where $\mathcal{B} = \mathcal{B}_1 \times \cdots \times \mathcal{B}_n$ and the following property is satisfied:

> If $G_1 \subseteq G_2 \subseteq H$, then for all $i = 1, \ldots, n$, $\mathcal{B}_i \dot{\cap} G_1 \subseteq \mathcal{B}_i \dot{\cap} G_2$.

Examples

1. For $i = 1, \ldots, n$ $\mathcal{B}_i \ := \ T_{-i}$ and for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \dot{\cap} G \ := \ S_{-i}$

Then $(\mathcal{B}, \dot{\cap})$ is the **pure** belief structure in H.

Examples

1. For $i = 1, \ldots, n$ $\mathcal{B}_i \; := \; T_{-i}$ and for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \,\dot{\cap}\, G \; := \; S_{-i}$

Then $(\mathcal{B}, \dot{\cap})$ is the **pure** belief structure in H.

2. Given a finite strategic game, let H be the mixed extension, so $H = (I_1, \ldots, I_n, U_1, \ldots, U_n)$ where $I_i = \Delta S_i$, where ΔX is the set of probability measures on X.

Then, for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \,\dot{\cap}\, G := \Pi_{j \neq i} \overline{S_j}$, where $\overline{S_j}$ is the convex hull of a set S_j of mixed strategies.

Examples

3. Assume H is a finite game. For $i = 1, \ldots, n$, $\mathcal{B}_i \; := \; \Pi_{j \neq i} \Delta T_j$ and for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \,\dot{\cap}\, G \; := \; \Pi_{j \neq i} \Delta S_j$

Examples

3. Assume H is a finite game. For $i = 1, \dots, n$, $\mathcal{B}_i \ := \ \Pi_{j \neq i} \Delta T_j$ and for a restriction $G = (S_1, \dots, S_n)$ of H, $\mathcal{B}_i \mathbin{\dot{\cap}} G \ := \ \Pi_{j \neq i} \Delta S_j$

4. Assume H is finite. For $i = 1, \dots, n$, $\mathcal{B}_i \ := \ \Pi_{j \neq i} \Delta T_{-i}$ and for a restriction $G = (S_1, \dots, S_n)$ of H, $\mathcal{B}_i \mathbin{\dot{\cap}} G \ := \ \Delta S_{-i}$

Examples

3. Assume H is a finite game. For $i = 1, \ldots, n$, $\mathcal{B}_i \ := \ \Pi_{j \neq i} \Delta T_j$ and for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \dot{\cap} G \ := \ \Pi_{j \neq i} \Delta S_j$

4. Assume H is finite. For $i = 1, \ldots, n$, $\mathcal{B}_i \ := \ \Pi_{j \neq i} \Delta T_{-i}$ and for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \dot{\cap} G \ := \ \Delta S_{-i}$

5. Assume H is a finite game. For $i = 1, \ldots, n$, $\mathcal{B}_i \ := \ \Pi_{j \neq i} \Delta^\circ T_j$, where for a set X, $\Delta^\circ X$ is the set of probabilities measures that assign positive probability to each element of X, and for a restriction $G = (S_1, \ldots, S_n)$ of H, $\mathcal{B}_i \dot{\cap} G \ := \ \Pi_{j \neq i} \Delta^\circ S_j$

Theorem. In all models where the players are *rational* and there is *common belief of rationality*, the players choose strategies that survive iterative removal of strictly dominated strategies (and, conversely...).

Subgames

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ be an *arbitrary* strategic game.

Subgames

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ be an *arbitrary* strategic game.

A **restriction** of H is a sequence $G = (G_1, \ldots, G_n)$ such that $G_i \subseteq H_i$ for all $i \in \{1, \ldots, n\}$.

The set of all restrictions of a game H ordered by componentwise set inclusion forms a complete lattice.

Game Models

Relational models: $\langle W, R_i \rangle$ where $R_i \subseteq W \times W$. Write $R_i(w) = \{v \mid wR_iv\}$.

Events: $E \subseteq W$

Knowledge/Belief: $\Box E = \{w \mid R_i(w) \subseteq E\}$

Common knowledge/belief:
$\Box^1 E = \Box E$
$\Box^{k+1} E = \Box\Box^k E$
$\Box^* E = \bigcap_{k=1}^{\infty} \Box^k E$

Fact. An event F is called **evident** provided $F \subseteq \Box F$. $w \in \Box^* E$ provided there is an evident event F such that $w \in F \subseteq \Box E$.

Game Models

Let $G = (G_1, \ldots, G_n)$ be a restriction of a game H.

A **knowledge/belief model of** G is a tuple $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ where $\langle W, R_1, \ldots, R_n \rangle$ is a knowledge/belief model and $\sigma_i : W \to G_i$.

Game Models

Let $G = (G_1, \ldots, G_n)$ be a restriction of a game H.

A **knowledge/belief model of** G is a tuple $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ where $\langle W, R_1, \ldots, R_n \rangle$ is a knowledge/belief model and $\sigma_i : W \rightarrow G_i$.

Given a model $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots \sigma_n \rangle$ for a restriction G and a sequence $\overline{E} = \{E_1, \ldots, E_n\}$ where $E_i \subseteq W$,

$$G_{\overline{E}} = (\sigma_1(E_1), \ldots, \sigma_n(E_n))$$

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- T is monotonic if for all G, G', $G \subseteq G'$ implies $T(G) \subseteq T(G')$

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- T is monotonic if for all G, G', $G \subseteq G'$ implies $T(G) \subseteq T(G')$
- G is a fixed-point if $T(G) = G$

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- T is monotonic if for all G, G', $G \subseteq G'$ implies $T(G) \subseteq T(G')$
- G is a fixed-point if $T(G) = G$
- νT is the largest fixed point of T

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- T is monotonic if for all G, G', $G \subseteq G'$ implies $T(G) \subseteq T(G')$
- G is a fixed-point if $T(G) = G$
- νT is the largest fixed point of T
- T^∞ is the "outcome of T: $T^0 = \top$, $T^{\alpha+1} = T(T^\alpha)$, $T^\beta = \bigcap_{\alpha<\beta} T^\alpha$, The outcome of iterating T is the least α such that $T^{\alpha+1} = T^\alpha$, denoted T^∞

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- T is monotonic if for all G, G', $G \subseteq G'$ implies $T(G) \subseteq T(G')$
- G is a fixed-point if $T(G) = G$
- νT is the largest fixed point of T
- T^∞ is the "outcome of T: $T^0 = \top$, $T^{\alpha+1} = T(T^\alpha)$, $T^\beta = \bigcap_{\alpha<\beta} T^\alpha$, The outcome of iterating T is the least α such that $T^{\alpha+1} = T^\alpha$, denoted T^∞
- **Tarski's Fixed-Point Theorem**: Every monotonic operator T has a (least and largest) fixed point $T^\infty = \nu T = \bigcup\{G \mid G \subseteq T(G)\}$.

Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- T is monotonic if for all G, G', $G \subseteq G'$ implies $T(G) \subseteq T(G')$
- G is a fixed-point if $T(G) = G$
- νT is the largest fixed point of T
- T^∞ is the "outcome of T: $T^0 = \top$, $T^{\alpha+1} = T(T^\alpha)$, $T^\beta = \bigcap_{\alpha<\beta} T^\alpha$, The outcome of iterating T is the least α such that $T^{\alpha+1} = T^\alpha$, denoted T^∞
- **Tarski's Fixed-Point Theorem**: Every monotonic operator T has a (least and largest) fixed point $T^\infty = \nu T = \bigcup\{G \mid G \subseteq T(G)\}$.
- T is contracting if $T(G) \subseteq G$. Every contracting operator has an outcome (T^∞ is well-defined)

Rationality Properties

$\varphi(s_i, G_i, G_{-i})$ holds between a strategy $s_i \in H_i$, a set of strategies G_i for player i and strategies G_{-i} of the opponents. Intuitively s_i is φ-optimal strategy for player i in the restricted game $\langle G_i, G_{-i}, u_1, \ldots, u_n \rangle$ (where the payoffs are suitably restricted).

Rationality Properties

$\varphi(s_i, G_i, G_{-i})$ holds between a strategy $s_i \in H_i$, a set of strategies G_i for player i and strategies G_{-i} of the opponents. Intuitively s_i is φ-optimal strategy for player i in the restricted game $\langle G_i, G_{-i}, u_1, \dots, u_n \rangle$ (where the payoffs are suitably restricted).

φ_i is **monotonic** if for all G_{-i}, $G'_{-i} \subseteq H_{-i}$ and $s_i \in H_i$

$$G_{-i} \subseteq G'_{-i} \text{ and } \varphi(s_i, H_i, G_{-i}) \text{ implies } \varphi(s_i, H_i, G'_{-i})$$

Removing Strategies

If $\varphi = (\varphi_1, \ldots, \varphi_n)$, then define $T_\varphi(G) = G'$ where

- $G = (G_1, \ldots, G_n)$, $G' = (G'_1, \ldots, G'_n)$,
- for all $i \in \{1, \ldots, n\}$, $\quad G'_i = \{s_i \in G_i \mid \varphi_i(s_i, H_i, G_{-i})\}$

Removing Strategies

If $\varphi = (\varphi_1, \ldots, \varphi_n)$, then define $T_\varphi(G) = G'$ where

- $G = (G_1, \ldots, G_n)$, $G' = (G'_1, \ldots, G'_n)$,
- for all $i \in \{1, \ldots, n\}$, $\quad G'_i = \{s_i \in G_i \mid \varphi_i(s_i, H_i, G_{-i})\}$

T_φ is contracting, so it has an outcome T^∞_φ

Removing Strategies

If $\varphi = (\varphi_1, \ldots, \varphi_n)$, then define $T_\varphi(G) = G'$ where

- $G = (G_1, \ldots, G_n)$, $G' = (G'_1, \ldots, G'_n)$,
- for all $i \in \{1, \ldots, n\}$, $\quad G'_i = \{s_i \in G_i \mid \varphi_i(s_i, H_i, G_{-i})\}$

T_φ is contracting, so it has an outcome T_φ^∞

If each φ_i is monotonic, then νT_φ exists and equals T_φ^∞.

Rational Play

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ a strategic game and $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ a model for H.

$\sigma_i(w)$ is the strategy player is using in state w.

$G_{R_i(w)}$ is a restriction of H giving i's view of the game.

Rational Play

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ a strategic game and $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ a model for H.

$\sigma_i(w)$ is the strategy player is using in state w.

$G_{R_i(w)}$ is a restriction of H giving i's view of the game.

Player i is φ_i-rational in the state w if $\varphi_i(\sigma_i(w), H_i, (G_{R_i(w)})_{-i})$ holds.

Rational Play

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ a strategic game and $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ a model for H.

$\sigma_i(w)$ is the strategy player is using in state w.

$G_{R_i(w)}$ is a restriction of H giving i's view of the game.

Player i is φ_i-rational in the state w if $\varphi_i(\sigma_i(w), H_i, (G_{R_i(w)})_{-i})$ holds.

$\mathbf{Rat}(\varphi) = \{w \in W \mid \text{ each player is } \varphi_i\text{-rational in } w\}$

$\Box\mathbf{Rat}(\varphi)$
$\Box^*\mathbf{Rat}(\varphi)$

Theorem (Apt and Zvesper).

- Suppose that each φ_i is monotonic. Then for all belief models for H,

$$G_{\mathbf{Rat}(\varphi)\cap B^*(\mathbf{Rat}(\varphi))} \subseteq T_\varphi^\infty$$

- Suppose that each φ_i is monotonic. Then for all knowledge models for H,

$$G_{K^*(\mathbf{Rat}(\varphi))} \subseteq T_\varphi^\infty$$

- For some standard knowledge model for H,

$$T_\varphi^\infty \subseteq G_{K^*(\mathbf{Rat}(\varphi))}$$

K. Apt and J. Zvesper. *The Role of Monotonicity in the Epistemic Analysis of Games*. Games, 1(4), pgs. 381-394, 2010.

Claim If each φ_i is monotonic, then $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)} \subseteq T_\varphi^\infty$.

Claim If each φ_i is monotonic, then $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)} \subseteq T^{\infty}_{\varphi}$.

Let s_i be an element of the ith component of $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)}$:
$s_i = \sigma_i(w)$ for some $w \in \mathbf{Rat}(\varphi) \cap \Box^*\mathbf{Rat}(\varphi)$

Claim If each φ_i is monotonic, then $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)} \subseteq T^\infty_\varphi$.

Let s_i be an element of the ith component of $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)}$:
$s_i = \sigma_i(w)$ for some $w \in \mathbf{Rat}(\varphi) \cap \Box^*\mathbf{Rat}(\varphi)$

there is an F such that $F \subseteq \Box F$ and

$$w \in F \subseteq \Box\mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$$

Claim If each φ_i is monotonic, then $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)} \subseteq T^\infty_\varphi$.

Let s_i be an element of the ith component of $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)}$:
$s_i = \sigma_i(w)$ for some $w \in \mathbf{Rat}(\varphi) \cap \Box^*\mathbf{Rat}(\varphi)$

there is an F such that $F \subseteq \Box F$ and

$$w \in F \subseteq \Box\mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$$

Claim. $G_{F\cap\mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

Claim If each φ_i is monotonic, then $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)} \subseteq T^\infty_\varphi$.

Let s_i be an element of the ith component of $G_{\mathbf{Rat}(\varphi)\cap\Box^*\mathbf{Rat}(\varphi)}$:
$s_i = \sigma_i(w)$ for some $w \in \mathbf{Rat}(\varphi) \cap \Box^*\mathbf{Rat}(\varphi)$

there is an F such that $F \subseteq \Box F$ and

$$w \in F \subseteq \Box\mathbf{Rat}(\varphi) = \{v \in W \mid \forall i \; R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$$

Claim. $G_{F\cap\mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

Since each φ_i is monotonic, T_φ is monotonic and by Tarski's fixed-point theorem, $G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T^\infty_\varphi$. But $s_i = \sigma_i(w)$ and $w \in F \cap \mathbf{Rat}(\varphi)$, so s_i is the ith component in T^∞_φ.

$F \subseteq \Box F$ and $w \in F \subseteq \Box \mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F\cap\mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

$F \subseteq \Box F$ and $w \in F \subseteq \Box \mathbf{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F \cap \mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F \cap \mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \mathbf{Rat}(\varphi)})$).

Let $w' \in F \cap \mathbf{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

$F \subseteq \Box F$ and $w \in F \subseteq \Box \mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F \cap \mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

Let $w' \in F \cap \mathbf{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \mathbf{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

$F \subseteq \Box F$ and $w \in F \subseteq \Box\mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F\cap\mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

Let $w' \in F \cap \mathbf{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \mathbf{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

F is evident, so $R_i(w') \subseteq F$. We also have $R_i(w') \subseteq \mathbf{Rat}(\varphi)$.

$F \subseteq \Box F$ and $w \in F \subseteq \Box \mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F\cap\mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

Let $w' \in F \cap \mathbf{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \mathbf{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

F is evident, so $R_i(w') \subseteq F$. We also have $R_i(w') \subseteq \mathbf{Rat}(\varphi)$.

Hence, $R_i(w') \subseteq F \cap \mathbf{Rat}(\varphi)$.

$F \subseteq \Box F$ and $w \in F \subseteq \Box \mathbf{Rat}(\varphi) = \{v \in W \mid \forall i \; R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F \cap \mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F \cap \mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \mathbf{Rat}(\varphi)})$).

Let $w' \in F \cap \mathbf{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \mathbf{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

F is evident, so $R_i(w') \subseteq F$. We also have $R_i(w') \subseteq \mathbf{Rat}(\varphi)$.

Hence, $R_i(w') \subseteq F \cap \mathbf{Rat}(\varphi)$.

This implies $(G_{R_i(w')}) \subseteq (G_{F \cap \mathbf{Rat}(\varphi)})_{-i}$, and so by monotonicity of φ_i, $\varphi_i(s_i, H_i, (G_{F \cap \mathbf{Rat}(\varphi)})_{-i})$ holds.

$F \subseteq \Box F$ and $w \in F \subseteq \Box \mathbf{Rat}(\varphi) = \{v \in W \mid \forall i\ R_i(v) \subseteq \mathbf{Rat}(\varphi)\}$

Claim. $G_{F\cap\mathbf{Rat}(\varphi)}$ is post-fixed point of T_φ ($G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$).

Let $w' \in F \cap \mathbf{Rat}(\varphi)$ and let $i \in \{1,\ldots,n\}$.

Since $w' \in \mathbf{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

F is evident, so $R_i(w') \subseteq F$. We also have $R_i(w') \subseteq \mathbf{Rat}(\varphi)$.

Hence, $R_i(w') \subseteq F \cap \mathbf{Rat}(\varphi)$.

This implies $(G_{R_i(w')}) \subseteq (G_{F\cap\mathbf{Rat}(\varphi)})_{-i}$, and so by monotonicity of φ_i, $\varphi_i(s_i, H_i, (G_{F\cap\mathbf{Rat}(\varphi)})_{-i})$ holds.

This means $G_{F\cap\mathbf{Rat}(\varphi)} \subseteq T_\varphi(G_{F\cap\mathbf{Rat}(\varphi)})$

$sd_i(s_i, G_i, G_{-i})$ is $\neg\exists s_i' \in G_i, \forall s_{-i} \in G_{-i} u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$

$sd_i(s_i, G_i, G_{-i})$ is $\neg\exists s_i' \in G_i, \forall s_{-i} \in G_{-i} u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$

$br_i(s_i, G_i, G_{-i})$ is $\exists\mu_i \in \mathcal{B}_i(G_{-i})\forall s_i' \in G_i, U_i(s_i, \mu_i) \geq U_i(s_i', \mu_i)$.

$sd_i(s_i, G_i, G_{-i})$ is $\neg\exists s_i' \in G_i, \forall s_{-i} \in G_{-i} u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$

$br_i(s_i, G_i, G_{-i})$ is $\exists \mu_i \in \mathcal{B}_i(G_{-i}) \forall s_i' \in G_i, U_i(s_i, \mu_i) \geq U_i(s_i', \mu_i)$.

$U_\varphi(G) = G'$ where $G_i' = \{s_i \in G_i \mid \varphi_i(s_i, G_i, G_{-i})\}$.

$sd_i(s_i, G_i, G_{-i})$ is $\neg\exists s_i' \in G_i, \forall s_{-i} \in G_{-i} u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$

$br_i(s_i, G_i, G_{-i})$ is $\exists \mu_i \in \mathcal{B}_i(G_{-i}) \forall s_i' \in G_i, U_i(s_i, \mu_i) \geq U_i(s_i', \mu_i)$.

$U_\varphi(G) = G'$ where $G_i' = \{s_i \in G_i \mid \varphi_i(s_i, G_i, G_{-i})\}$.

Note: U_φ is *not* monotonic.

Corollary. For all belief models, $G_{\mathbf{Rat}(br)\cap\Box^*\mathbf{Rat}(br)} \subseteq U_{sd}^{\infty}$. For all G, we have

$$T_{br}(G) \subseteq T_{sd}(G)$$

$$T_{sd}(G) \subseteq U_{sd}(G)$$

Then, $T_{sd}^{\infty} \subseteq U_{sd}^{\infty}$.

Corollary. For all belief models, $G_{\mathbf{Rat}(br)\cap\Box^*\mathbf{Rat}(br)} \subseteq U^\infty_{sd}$. For all G, we have

$$T_{br}(G) \subseteq T_{sd}(G)$$

$$T_{sd}(G) \subseteq U_{sd}(G)$$

Then, $T^\infty_{sd} \subseteq U^\infty_{sd}$.

Fact. Consider two operators T_1, T_2 on $(D, \subseteq)$ such that,

- for all G, $T_1(G) \subseteq T_2(G)$
- T_1 is monotonic
- T_2 is contracting

Then, $T^\infty_1 \subseteq T^\infty_2$.

This analysis does not work for weak dominance...

Rationality

Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game and $\mathcal{T} = \langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ a type space for G.

For each $t_i \in T_i$, we can define a probability measure $p_{t_i} \in \Delta(S_{-i})$:

$$p_{t_i}(s_{-i}) = \sum_{t_{-i} \in T_{-i}} \lambda_i(t_i)(s_{-i}, t_{-i})$$

Rationality and common belief of rationality (RCBR) in the matrix

IESDS

			2	
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	m	0, 4	0, 0	4, 0

IESDS

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	m	0, 4	0, 0	4, 0

$\rightarrowtail$

		2	
		l	c
1	t	3, 3	1, 1
	m	1,1	3, 3
	b	0, 4	0, 0

IESDS

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	m	0, 4	0, 0	4, 0

$\rightarrowtail$

		2	
		l	c
1	t	3, 3	1, 1
	m	1,1	3, 3
	b	0, 4	0, 0

$\rightarrowtail$

		2	
		l	c
1	t	3, 3	1, 1
	m	1,1	3, 3

1's types

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

2's types

$\lambda_2(s_1)$

	t	m	b
t_1	0.5	0.5	0
t_2	0	0	0

$\lambda_2(s_2)$

	t	m	b
t_1	0.25	0.25	0
t_2	0.25	0.25	0

$\lambda_2(s_3)$

	t	m	b
t_1	0.5	0	0
t_2	0	0	0.5

2

1

	l	c	r
t	3, 3	1, 1	0, 0
m	1,1	3, 3	1, 0
b	0, 4	0, 0	4, 0

$\lambda_2(s_1)$

	t	m	b
t_1	0.5	0.5	0
t_2	0	0	0

$\lambda_2(s_2)$

	t	m	b
t_1	0.25	0.25	0
t_2	0.25	0.25	0

$\lambda_2(s_3)$

	t	m	b
t_1	0.5	0	0
t_2	0	0	0.5

2

		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

		t	m	b
$\lambda_2(s_2)$	t_1	0.25	0.25	0
	t_2	0.25	0.25	0

		t	m	b
$\lambda_2(s_3)$	t_1	0.5	0	0
	t_2	0	0	0.5

- l and c are rational for both s_1 and s_2.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

		t	m	b
$\lambda_2(s_2)$	t_1	0.25	0.25	0
	t_2	0.25	0.25	0

		t	m	b
$\lambda_2(s_3)$	t_1	0.5	0	0
	t_2	0	0	0.5

- l and c are rational for both s_1 and s_2.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

$\lambda_2(s_1)$

	t	m	b
t_1	0.5	0.5	0
t_2	0	0	0

$\lambda_2(s_2)$

	t	m	b
t_1	0.25	0.25	0
t_2	0.25	0.25	0

$\lambda_2(s_3)$

	t	m	b
t_1	0.5	0	0
t_2	0	0	0.5

- l and c are rational for both s_1 and s_2.
- l is the only rational action for s_3.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

		t	m	b
$\lambda_2(s_2)$	t_1	0.25	0.25	0
	t_2	0.25	0.25	0

		t	m	b
$\lambda_2(s_3)$	t_1	0.5	0	0
	t_2	0	0	0.5

- l and c are rational for both s_1 and s_2.
- l is the only rational action for s_3.
- Whatever her type, it is never rational to play r for 2.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		l	c	r
$\lambda_1(t_2)$	s_1	0	0.5	0
	s_2	0	0	0.5
	s_3	0	0	0

- t and m are rational for t_1.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		l	c	r
$\lambda_1(t_2)$	s_1	0	0.5	0
	s_2	0	0	0.5
	s_3	0	0	0

- t and m are rational for t_1.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		l	c	r
$\lambda_1(t_2)$	s_1	0	0.5	0
	s_2	0	0	0.5
	s_3	0	0	0

- t and m are rational for t_1.
- m and b are rational for t_2.

$\lambda_2(s_1)$

	t	m	b
t_1	0.5	0.5	0
t_2	0	0	0

$\lambda_2(s_2)$

	t	m	b
t_1	0.25	0.25	0
t_2	0.25	0.25	0

$\lambda_2(s_3)$

	t	m	b
t_1	0.5	0	0
t_2	0	0	0.5

$\lambda_2(s_1)$

	t	m	b
t_1	0.5	0.5	0
t_2	0	0	0

$\lambda_2(s_2)$

	t	m	b
t_1	0.25	0.25	0
t_2	0.25	0.25	0

$\lambda_2(s_3)$

	t	m	b
t_1	0.5	0	0
t_2	0	0	0.5

- All of 2's types believe that 1 is rational.

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

- Type t_1 of 1 believes that 2 is rational.

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

- Type t_1 of 1 believes that 2 is rational.
- But type t_2 doesn't! (1/2 probability that 2 is playing r.)

$\lambda_2(s_1)$

	t	m	b
t_1	0.5	0.5	0
t_2	0	0	0

$\lambda_2(s_2)$

	t	m	b
t_1	0.25	0.25	0
t_2	0.25	0.25	0

$\lambda_2(s_3)$

	t	m	b
t_1	0.5	0	0
t_2	0	0	0.5

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

		t	m	b
$\lambda_2(s_2)$	t_1	0.25	0.25	0
	t_2	0.25	0.25	0

		t	m	b
$\lambda_2(s_3)$	t_1	0.5	0	0
	t_2	0	0	0.5

- Only type s_1 of 2 believes that 1 is rational and that 1 believes that 2 is also rational.

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

$\lambda_1(t_1)$

	l	c	r
s_1	0.5	0.5	0
s_2	0	0	0
s_3	0	0	0

$\lambda_1(t_2)$

	l	c	r
s_1	0	0.5	0
s_2	0	0	0.5
s_3	0	0	0

- Type t_1 of 1 believes that 2 is rational and that 2 believes that 1 believes that 2 is rational.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

- No further iteration of mutual belief in rationality eliminate some types or strategies.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

- No further iteration of mutual belief in rationality eliminate some types or strategies.
- So at all the states in $\{(t_1, s_1)\} \times \{t, m\} \times \{l, c\}$ we have rationality and common belief in rationality.

		2		
		l	c	r
1	t	3, 3	1, 1	0, 0
	m	1,1	3, 3	1, 0
	b	0, 4	0, 0	4, 0

		l	c	r
$\lambda_1(t_1)$	s_1	0.5	0.5	0
	s_2	0	0	0
	s_3	0	0	0

		t	m	b
$\lambda_2(s_1)$	t_1	0.5	0.5	0
	t_2	0	0	0

- No further iteration of mutual belief in rationality eliminate some types or strategies.
- So at all the states in $\{(t_1, s_1)\} \times \{t, m\} \times \{l, c\}$ we have rationality and common belief in rationality.
- But observe that $\{t, m\} \times \{l, c\}$ is precisely the set of profiles that survive IESDS.

RCBR

Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N}\rangle$ be a strategic game and $\mathcal{T} = \langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S\rangle$ a type space for G.

The set of states (pairs of strategy profiles and type profiles) where player i chooses **rationally** is:

$$\mathsf{Rat_i} \ := \ \{(s_i, t_i) \mid s_i \text{ is a best response to } p_{t_i}\}$$

The event that all players are *rational* is
$\mathsf{Rat} = \{(s, t) \mid \text{ for all } i, (s_i, t_i) \in \mathsf{Rat_i}\}$.

RCBR

A type $t_i \in T_i$ **believes** an event $E_{-i} \subseteq S_{-i} \times T_{-i}$ if $\lambda_i(t_i)(E_{-i}) = 1$; let $B_i(E_{-i}) = \{(s_i, t_i) \mid t_i \text{ believes } E_{-i}\}$.

$R_i^1 = \mathrm{Rat}_i$,

for $m \geq 1$, $R_i^{m+1} = R_i^m \cap B_i(R_{-i}^m)$

$$RCBR_i = \bigcap_{m \geq 1} R_i^m \text{ and } RCBR = \Pi_{i \in N} RCBR_i$$

BRS

Let $S_i^0 = S_i$ for all $i \in N$. For $m \geq 0$, let S_i^{m+1} be the set of strategies that are best replies to conjectures $\mu_{-i} \in \Delta S_{-i}^m$. The set $S_i^\infty = \bigcap_{m\geq 0} S_i^m$ is the set of (correlated) rationalizable strategies of Player i.

A set $B = \Pi_{i\in N} B_i \subseteq S = \Pi_{i\in N} S_i$ is a **best-reply set** (or BRS) if, for all players $i \in N$, every $s_i \in B_i$ is a best reply to a belief $\mu_{-i} \in \Delta B_{-i}$. B is a **full BRS** if, for every $s_i \in B_i$, there is a belief $\mu_{-i} \in \Delta B_{-i}$ that rationalizes s_i and such that all best replies to μ_{-i} are also in B_i.

Theorem (Brandenburger and Dekel, Tan and da Costa Werlang) Fix a game $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$.

1. In any type structure $\langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ for G, $\text{proj}_S RCBR$ is a full BRS.
2. In any *complete* type structure $\langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ for G, $\text{proj}_S RCBR = S^\infty$.
3. For every full BRS B, there exists a finite type structure $\langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ for G such that $\text{proj}_S RCBR = B$.

Theorem (Brandenburger and Dekel, Tan and da Costa Werlang) Fix a game $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$.

1. In any type structure $\langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ for G, $\text{proj}_S RCBR$ is a full BRS.
2. In any *complete* type structure $\langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ for G, $\text{proj}_S RCBR = S^\infty$.
3. For every full BRS B, there exists a finite type structure $\langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ for G such that $\text{proj}_S RCBR = B$.

			2	
		l	c	r
1	t	4,4	1,1	0,0
	m	1,1	5,5	0,0
	d	0,1	0,1	6,0

	b
l	1
c	0
r	0

	a
t	1
m	0
d	0

A. Friedenberg and J. Kiesler. *Iterated Dominance Revisited*. Working paper, 2011.

2

		l	c	r
1	t	4,4	1,1	0,0
	m	1,1	5,5	0,0
	d	0,1	0,1	6,0

	b
l	1
c	0
r	0

	a
t	1
m	0
d	0

- The projection of $RCBR$ is $\{(t, l)\}$

A. Friedenberg and J. Kiesler. *Iterated Dominance Revisited*. Working paper, 2011.

			2	
		l	c	r
1	t	4,4	1,1	0,0
	m	1,1	5,5	0,0
	d	0,1	0,1	6,0

	b
l	1
c	0
r	0

	a
t	1
m	0
d	0

- The projection of $RCBR$ is $\{(t, l)\}$
- This is not the entire *ISDS* set

A. Friedenberg and J. Kiesler. *Iterated Dominance Revisited*. Working paper, 2011.

2

		l	c	r
1	t	4,4	1,1	0,0
	m	1,1	5,5	0,0
	d	0,1	0,1	6,0

	b
l	1
c	0
r	0

	a
t	1
m	0
d	0

- The projection of *RCBR* is $\{(t, l)\}$
- This is not the entire *ISDS* set
- "Game independent" conditions and *rich* type structures

A. Friedenberg and J. Kiesler. *Iterated Dominance Revisited*. Working paper, 2011.