Admissibility in Games

Adam Brandenburger, Amanda Friedenberg,
H. Jerome Keisler

January 14, 2007

“Admissibility in Games,” by Adam Brandenburger, Amanda Friedenberg, and H. Jerome Keisler

“The Power of Paradox: Some Recent Developments in Interactive Epistemology,” by Adam Brandenburger, forthcoming in *International Journal of Game Theory*

Both available at www.stern.nyu.edu/~abranden

The Question

What is the implication of supposing in a game that

each player is rational

each player thinks the other players are rational

and so on?

"We wish to find the mathematically complete principles which define 'rational behavior' for the participants.… The rules of rational behavior must provide definitely for the possibility of irrational conduct on the part of others"

--*Theory of Games and Economic Behavior*, by John von Neumann and Oskar Morgenstern, 1944

A First Answer

The players choose iteratively undominated strategies:

delete from the game all (strongly) dominated strategies

delete from the remaining game all dominated strategies

and so on

But what if rationality is taken to include an **admissibility** requirement—i.e., the avoidance of weakly dominated strategies?

Why Admissibility?

Iterated admissibility (**IA**) gives:

- sharp answers in many games of applied interest—auctions, voting games, Bertrand
- the backward-induction outcomes in perfect-information trees (under a payoff condition)
- the forward-induction outcome in well-known signalling games*

Admissibility is prima facie reasonable: It says that a player takes into consideration all strategies for the other players

Admissibility is decision-theoretically equivalent to invariance à la Dalkey (1953)-Thompson (1952)

*E.g. the original signalling example of Kohlberg-Mertens (1986); the Burn-a-Dollar game of van Damme (1989) and Ben Porath-Dekel (1992); the Beer-Quiche game of Cho-Kreps (1987) (viewed as a two-player game)

The Basic Challenge

Recall: A strategy is admissible iff there is a strictly positive probability measure on the strategy profiles of the other players, under which it is optimal

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

Ann should assign positive probability to both *L* and *R*

Bob should assign positive probability to both *U* and *D*

Then Bob will play *L*

But then Ann should assign probability 1 to *L*?

See Samuelson (1992)

Lexicographic Probabilities

Allow Ann at the same time both to include and to exclude a strategy of Bob's

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

Ann has a **lexicographic probability system** (Blume-Brandenburger-Dekel 1991):

Her primary hypothesis assigns probability 1 to *L*

Her secondary hypothesis assigns probability 1 to *R*

We'll say Ann **assumes** Bob is rational

Similarly for Bob

Rationality and Common Assumption of Rationality

Consider the sequence of conditions:

(a1) Ann is rational

(a2) Ann is rational and assumes (b1)

(a3) Ann is rational and assumes (b1)
and assumes (b2)

etc.

(b1) Bob is rational

(b2) Bob is rational and assumes (a1)

(b3) Bob is rational and assumes (a1)
and assumes (a2)

etc.

Question: What strategies can be played under **RCAR**?

Type Structures

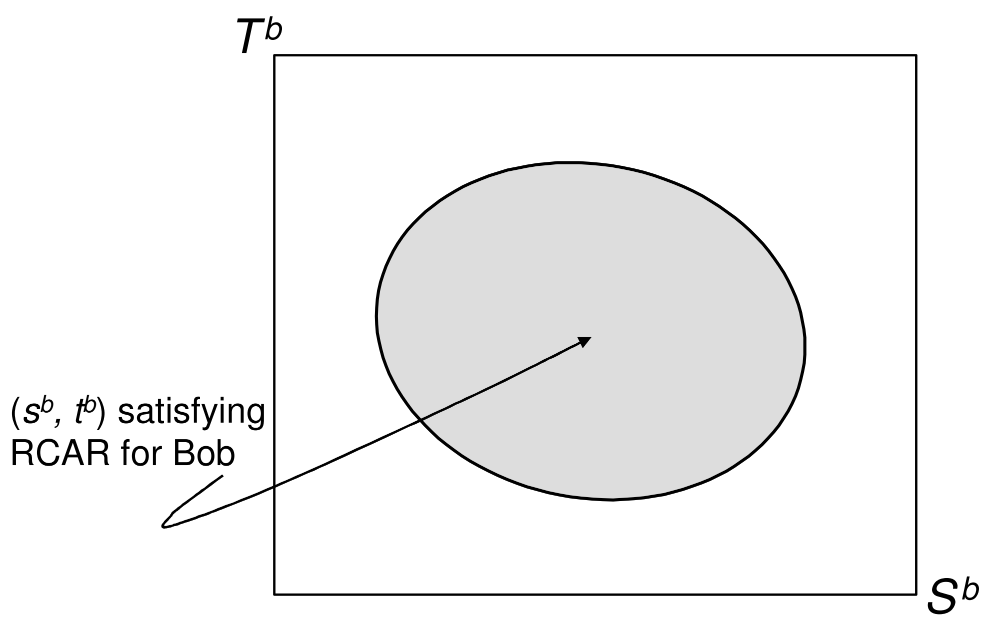
Let T^a, T^b be spaces of types for Ann and Bob resp.

Each type t^a for Ann is associated with an LPS on $S^b \times T^b$, and likewise for Bob

A state of the world is a 4-tuple (s^a, t^a, s^b, t^b)

A pair (s^a, t^a) is **rational** if:

- the LPS associated with t^a has full support
- strategy s^a lex-maximizes Ann's expected payoff under this LPS

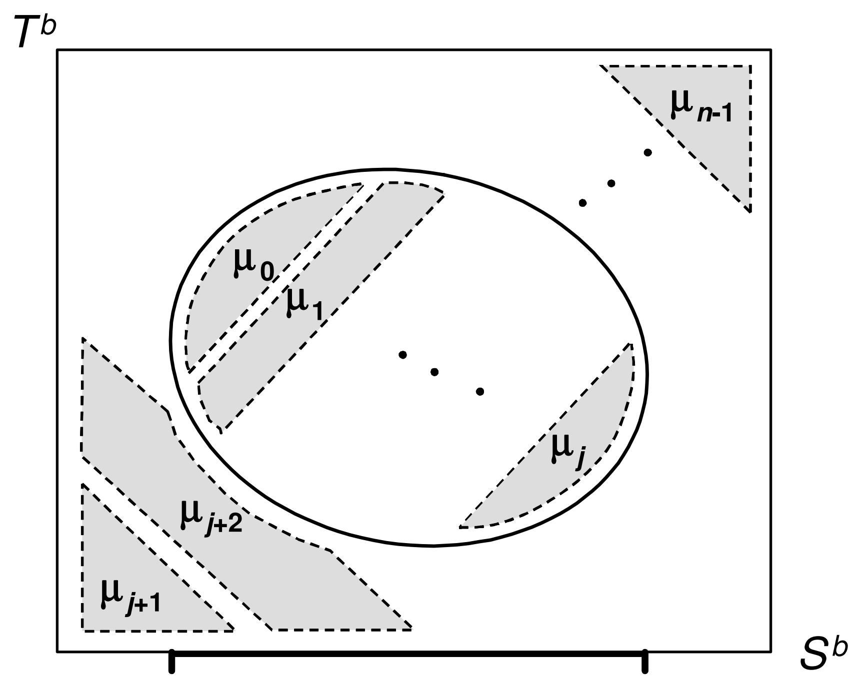


Characterizing RCAR

Fix a strategy-type pair (s^a, t^a) that satisfies RCAR for Ann

Let $(\mu_0, \ldots, \mu_{n-1})$ be the LPS associated with t^a

By a conjunction property for assumption, t^a assumes RCAR for Bob:



The Candidate

Take the set of all states (s^a, t^a, s^b, t^b) satisfying RCAR

Let $Q^a \times Q^b$ be its projection into $S^a \times S^b$

Then we see that $Q^a \times Q^b$ has the two properties:

(i) each $s^a \in Q^a$ is admissible (i.e. admissible wrt S^b)

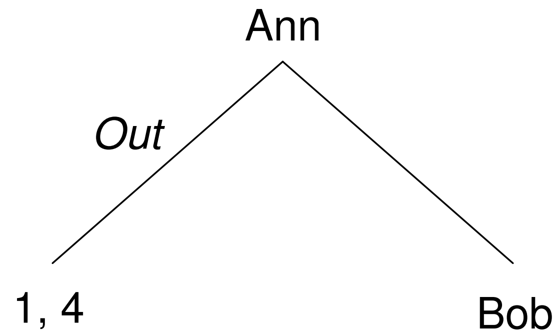
(ii) each $s^a \in Q^a$ is admissible wrt Q^b

and likewise with *a* and *b* interchanged

Note the similarity to a Pearce (1984) best-response set

But these two properties do not characterize RCAR

Convex Combinations



Ann		L	R
	U	-1, 3	-1, 0
Ann	M	2, 0	0, 3
	D	0, 0	2, 3

$\{Out\} \times \{L, R\}$ has properties (i) and (ii)

But *Out* cannot be played under RCAR!

Self-Admissible Set

An **SAS** is a subset $Q^a \times Q^b \subseteq S^a \times S^b$ satisfying:

(i) each $s^a \in Q^a$ is admissible (i.e. admissible wrt S^b)

(ii) each $s^a \in Q^a$ is admissible wrt Q^b

(iii) if $s^a \in Q^a$, and r^a is part of a convex combination of strategies for Ann that is equivalent for her to s^a, then $r^a \in Q^a$

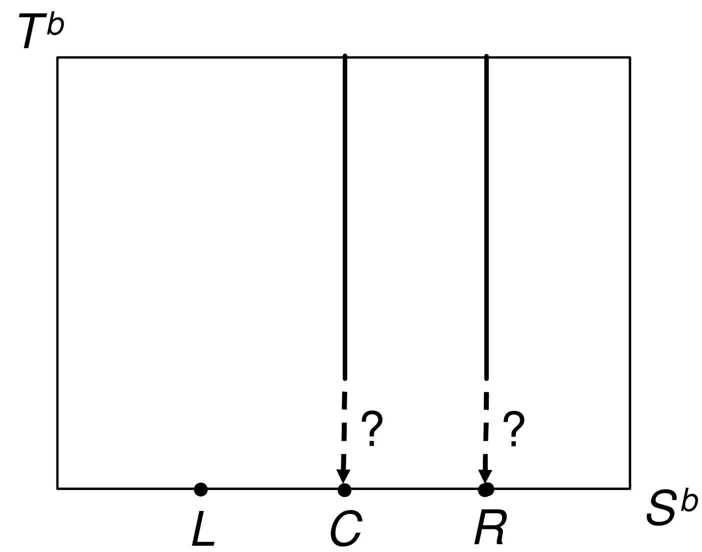
and likewise with *a* and *b* interchanged

But we still don't have a characterization of RCAR

Irrationality

Bob

		L	C	R
	U	4, 0	4, 1	0, 1
Ann	M	0, 0	0, 1	4, 1
	D	3, 0	2, 1	2, 1



$\{U, M, D\} \times \{C, R\}$ is an SAS

Can D be played under RCAR?

Conceptually: If Ann considers everything possible, she should, in particular, take into account the possibility that Bob doesn't consider everything possible

Example kindly provided by Pierpaolo Battigalli

Characterization Theorem

Start with a game and an associated type structure. Let $Q^a \times Q^b$ be the projection into $S^a \times S^b$ of the states (s^a, t^a, s^b, t^b) satisfying RCAR. Then $Q^a \times Q^b$ is an SAS of the game.

Conversely, start with a game and an SAS $Q^a \times Q^b$. There is a type structure (with non-full-support types) such that $Q^a \times Q^b$ is the projection into $S^a \times S^b$ of the states (s^a, t^a, s^b, t^b) satisfying RCAR.

SAS's have 'good' properties—yielding:

Existence (the IA set is an SAS)

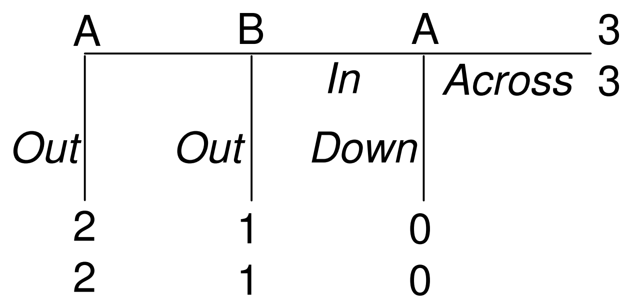
Invariance à la Kohlberg-Mertens (1986)

Projection à la Kohlberg-Mertens (1986)

Nash outcomes in perfect-information games satisfying a payoff condition

Iterated Admissibility

What epistemic conditions select the IA set, from the family of SAS's?



LPS associated with t^a:

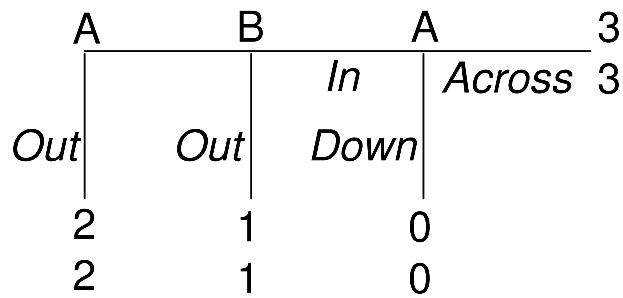
	Out	In
T^b t^b	1	[1]
	S^b	

LPS associated with t^b:

	Out	Down	Across
T^a t^a	1	[1]	[[1]]
		S^a	

The RCAR set is {(*Out*, t^a, *Out*, t^b)}

Completeness



LPS associated with t^a:

T^b		*Out*	*In*
	t^b	1	[1]

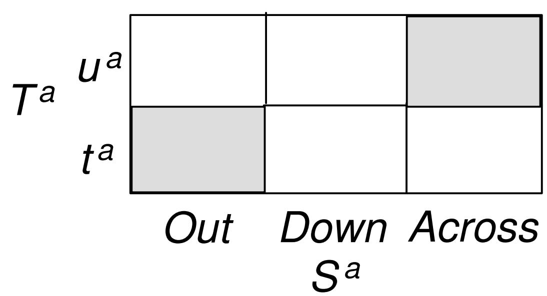
S^b

LPS associated with u^a:

T^b		*Out*	*In*
	t^b	[1]	1

S^b

LPS associated with t^b:



Now (*Out*, t^a, *Out*, t^b) does not satisfy RCAR!

Second Characterization Theorem

Call a type structure **complete** if

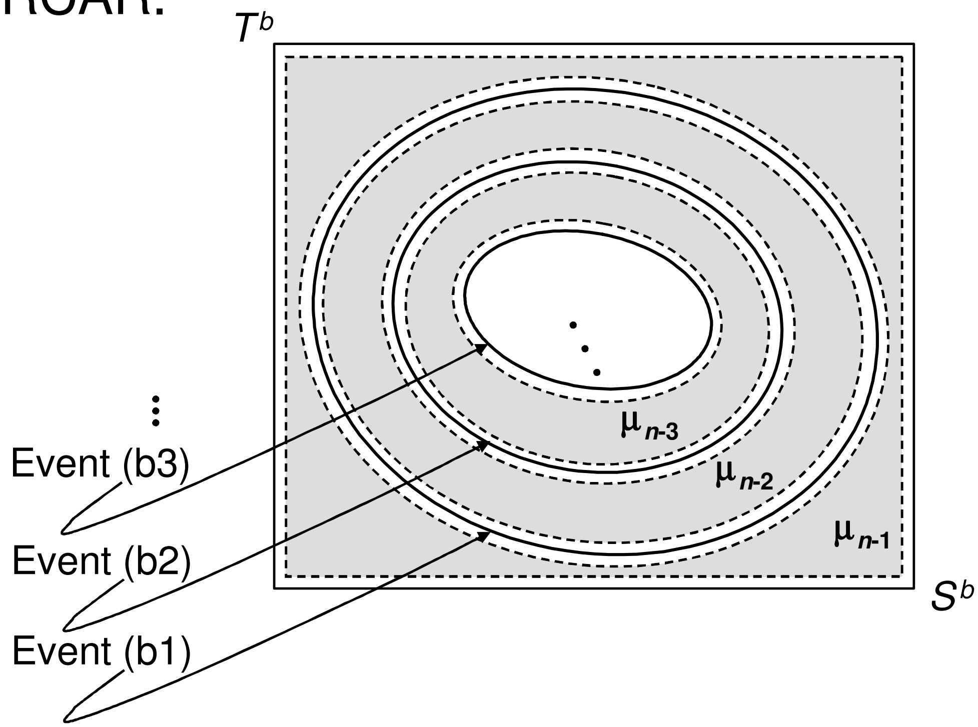
> the range of the map from T^a (Ann's type space) to the space of LPS's on $S^b \times T^b$ (Bob's strategy space cross Bob's type space) properly contains the set of full-support LPS's on $S^b \times T^b$
>
> and likewise with *a* and *b* interchanged

Complete type structures exist for every finite game

Start with a game and an associated complete type structure. Let $Q^a \times Q^b$ be the projection into $S^a \times S^b$ of the states (s^a, t^a, s^b, t^b) satisfying rationality and mth-order assumption of rationality. Then $Q^a \times Q^b$ is the set of strategies that survive $(m+1)$ rounds of IA.

An Impossibility Theorem

Start with a game in which Ann has more than one “strategically distinct” strategy, and an associated continuous complete type structure. Then no state satisfies RCAR.



Let t^a be a type for Ann that assumes each of (b1), (b2), …

Let $(\mu_0, \ldots, \mu_{n-1})$ be the LPS associated with t^a

Discussion

It seems there is a theoretical limit on players' ability to reason about all possibilities in a game

admissibility asks players to take all states into consideration

RCAR asks players to assumes "rationality and *m*th-order assumption of rationality" for all *m*

completeness asks players to consider all possible types that are implied by the model

"It is as if every time we think we finally get a hold on what rational behaviour means, we find ourselves having grasped only a shadow. Maybe this means there is excessive $\acute{\upsilon}\beta\rho\iota\varsigma$ *in this endeavour: that rationality is something belonging to the gods themselves, and that should not be stolen from them. Maybe it is the tree of knowledge itself, that we should not touch?"*
--Mertens (1989)

Or, perhaps, we are allowed to know that rationality in its 'ultimate form' cannot be