Decisions, Games & Logic '08

Amsterdam, July 1st 2008

Logical Foundations of Game Theory

Lecture 1: Static Games

Oliver Board, University of Pittsburgh

What is the correct solution to the following games?

	L	*R*
U	1, 1	-1, -1
D	0, 3	0, 3

Rationalizable outcomes: (*U*, *L*), (*U*, *R*), (*D*, *L*), (*D*, *R*)

Nash equilibria: (*U*, *L*), (*D*, *R*)

Trembling-hand perfect equilibrium: (*U*, *L*)

	L	*R*
U	10, 10	0, 0
D	0, 0	1, 1

Rationalizable outcomes: (*U*, *L*), (*U*, *R*), (*D*, *L*), (*D*, *R*)

Nash equilibria: (*U*, *L*), (*D*, *R*), $\left(\frac{1}{11}U \;\&\; \frac{10}{11}D, \frac{1}{11}L \;\&\; \frac{10}{11}R\right)$

Pareto-dominant Nash equilibrium: (*U*, *L*)

	L	*R*
A	6, 200	-10000, 0
B	5, 3	-1000, 5
C	3, 4	3, 6
D	-1000, 0	20, 3

Nash equilibria: (*A*, *L*), (*D*, *R*), mixed

Common sense: (*C*, *L*), (*C*, *R*)

An alternative approach

~~What is the right solution concept for games?~~

When is a given solution concept appropriate?

i.e. can we find conditions on the players' beliefs about each other which guarantee that they will behave according to some solution concept?

Alternatively, given some set of assumptions about the players' beliefs, what will they end up doing?

A model of beliefs about beliefs

An *information structure* for *n* agents:

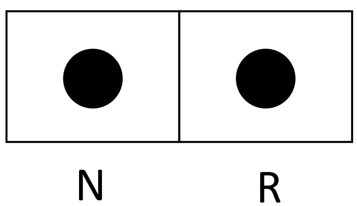
$$\langle W, \mathcal{I}_1, \dots, \mathcal{I}_n \rangle$$

- W is a set of *possible worlds* or *states*
- $\mathcal{I}_i : W \to 2^W$ is agent i's information function

$\mathcal{I}_i(w)$ should be thought of as the states agent i considers possible when the true state is *w*

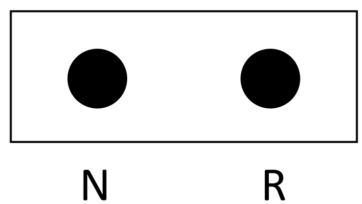
For example, consider the recent match between the Netherlands and Romania…

Adam is at the game



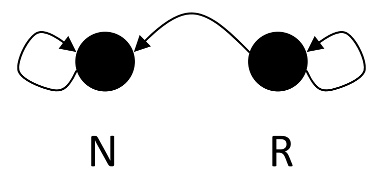
$$\mathcal{J}_A(N) = N, \mathcal{J}_A(R) = R$$

Bob is hiking in the Alps



$$\mathcal{J}_B(N) = \mathcal{J}_B(R) = \{N, R\}$$

Carl is in his office in Amsterdam



$$\mathcal{J}_C(N) = N, \mathcal{J}_C(R) = \{N, R\}$$

The Belief Operator

For any event $E \subseteq W$, the event that agent i believes E is given by:

$$B_i(E) = \{w \mid \mathcal{I}_i(w) \subseteq E\}$$

Intuitively, the agent believes something if it holds at every state she considers possible.

$B_i(E)$ satisfies the following three properties:

B1 $B_i(W) = W$

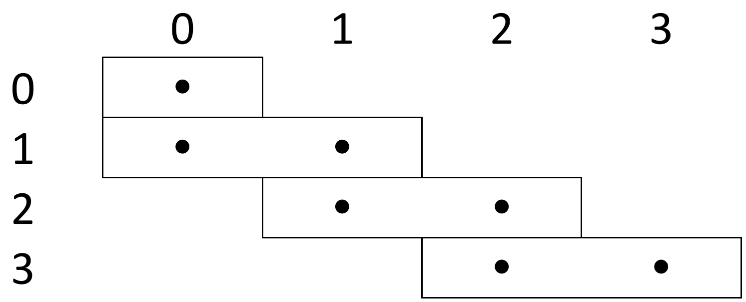
B2 if $E \subseteq F$ then $B_i(E) \subseteq B_i(F)$

B3 $B_i(E) \cap B_i(F) = B_i(E \cap F)$

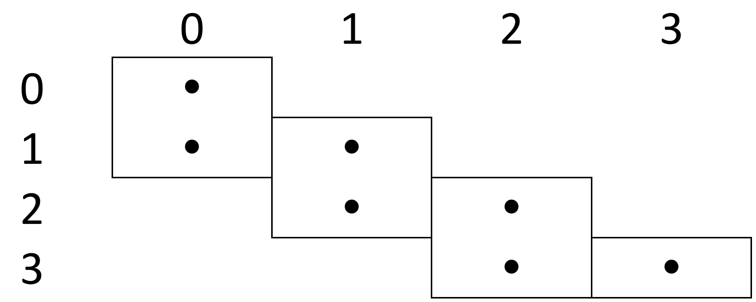
Example

A father wants to share a small sum of money between his two kids, Alice and Bob.

Alice is older, so if there is an odd number of dollars, she gets $1 more than Bob.



$\mathcal{J}_A$



$\mathcal{J}_B$

Consider the event $E = \{00,10,11,21\}$ ("Bob has less than $2")

$$E = \{00,10,11,21\}$$

$$B_A E = \{00,10,11\}$$

$$B_B B_A E = \{00,10\}$$

$$B_A B_B B_A E = \{00\}$$

$$B_B B_A B_B B_A E = \emptyset$$

Partitional Information Structures

Economists usually impose two additional constraints on the information functions:

I1 $w \in \mathcal{I}_i(w)$ for all $w \in W$

I2 if $x \in \mathcal{I}_i(w)$, then $\mathcal{I}_i(w) = \mathcal{I}_i(x)$

If $\mathcal{I}_i$ satisfies I1 and I2, then it partitions the state space, and the belief operator satisfies:

B4 $\mathrm{B_i E \subseteq E}$

B5 $B_i E \subseteq B_i B_i E$

B6 $\sim B_i E \subseteq B_i \sim B_i E$ (where $\sim X = W \setminus X$)

Henceforth, we assume that each $\mathcal{I}_i$ satisfies I1 and I2, and talk about knowledge (K_i) rather than belief

Common Knowledge

There is *common knowledge* of an event if everyone knows it, everyone knows that everyone knows it, and so on.

We define a common knowledge operator, $C : 2^W \to 2^W$ as follows

- First, define the operator $K : 2^W \to 2^W$ ("everyone knows that…")

$$K(E) = \cap_{i=1}^{n} K_i(E)$$

- Next, iterate the K operator

$$K^1(E) = K(E)$$

$$K^{k+1}(E) = K^k\big(K(E)\big) \text{ for } k \geq 1$$

- Finally, define C

$$C(E) = \cap_{k=1}^{\infty} K^k(E)$$

Characterizing Common Belief

An event $F \subseteq W$ is *self-evident* to agent i if $F = K_i(F)$

The *meet* of a collection of partitions

$$\mathcal{I} = \mathcal{I}_1 \wedge \mathcal{I}_2 \wedge \ldots \wedge \mathcal{I}_n$$

is the finest partition which is at least as course as each $\mathcal{I}_i$

1. Event E is common knowledge at state w if and only if there is an event F which is self-evident to every agent, such that $w \in F \subseteq E$

2. Event E is common knowledge at state w if and only if $\mathcal{I}(w) \subseteq E$, i.e.

$$C(E) = \{w \mid \mathcal{I}(w) \subseteq E\}$$

Introducing Probabilities

Augment the information structure with a probability measure over W for each agent

$$\langle W, \mathcal{I}_1, \ldots, \mathcal{I}_n, p_1, \ldots, p_n \rangle$$

Agent i's prior beliefs about E are computed by summing over the states in E

$$p_i(E) = \sum_{w \in E} p_i(W)$$

Posterior beliefs at a given state are obtained by conditionalization

$$p_i^w(E) = \frac{p_i(\mathcal{I}_1(w) \cap E)}{p_i\big(\mathcal{I}_1(w)\big)}$$

Example

	0	1	2	3
0	1/2			
1	1/4	1/8		
2		1/16	1/32	
3			1/64	1/64

p_A

Recall event $E = \{00,10,11,21\}$ ("Bob has less than \$2")

$$p_A(E) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}$$

$$p_A^{21}(E) = \frac{p_A(21)}{p_A(21)+p_A(22)} = \frac{1/16}{1/16+1/32} = \frac{2}{3}$$

Games

A *strategic form game* is defined as

$$G = \langle n, S_1, \dots, S_n, u_1, \dots, u_n \rangle,$$

where

- n is a set of players
- S_i is player i's strategy set
- $u_i : S_1 \times \dots \times S_n \to \mathbb{R}$ is player i's payoff function

Useful notation:

- $S = S_1 \times \dots \times S_n$ (the set of strategy profiles)
- $S_{-i} = \times_{j \neq i} S_j$ (the set of strategy profiles of i's opponents
- $\Sigma_\mathrm{i} = \Delta(S_i)$ (the set of mixed strategies for player i)
- $\Sigma = \Sigma_1 \times \dots \times \Sigma_n$ (the set of mixed strategy profiles – note: no correlation allowed)

Models of Games

A *model* for a game *G* consists of

$$\langle W, \mathcal{I}_1, \dots, \mathcal{I}_n, p_1, \dots, p_n, f_1, \dots, f_n \rangle$$

where

- $W, \mathcal{I}_i, \text{and } p_i$ are as before
- $f_i : W \to S_i$ is player i's *strategy function*

f_i describes what strategy player i will choose at each state

We assume: $f_i(w) = f_i(x)$ for all $x \in \mathcal{I}_i(w)$

(each player knows what she is going to do)

Expected Utility

The expected utility for player i if she plays s_i at state w is given by

$$EU_i^w(s_i) = \sum_{x \in W} p_i^w(x) \cdot u_i\big(s_i, f_{-i}(x)\big).$$

Note:

1. $p_i^w(x)$ will typically be equal to zero for many states x
2. This expression is not player i's *actual* expected utility, unless $s_i = f_i(w)$

We are making an implicit *causal independence* assumption: if player i were to play s_i instead of $f_i(w)$, her opponents would not change what they are doing.

Rationality

A player is *rational* if her expected utility given what she's actually doing is at least as high as it would be if she did something else instead.

Formally:

$$Rat_i = \left\{w \mid EU_i^w\big(f_i(w)\big) \geq EU_i^w(s_i), \text{for all } s_i \in S_i\right\}$$

$$Rat = \bigcap_{i=1}^{n} Rat_i$$

$$CKR = C(Rat)$$

Example

	Pub	Café
Pub	3, 1	0, 0
Café	0, 0	1, 3

G

	Pub	Café
Pub	•	•
Café	•	•

$\mathcal{I}_1$

	Pub	Café
Pub	1/3	1/6
Café	1/3	1/6

p_1

At every state, player 1 believes:

- with probability 2/3 that player 2 plays “Pub”
- with probability 1/3 that player 2 plays “Café”

So, $Rat_1 = \{PP, PC\}$

Note: at PC, player 1 is rational even though he gets a bad outcome; at CC, he is not rational even though he gets a good outcome.

Strictly Dominated Strategies

A strategy s_i is *strictly dominated* in G if there is some $\sigma_i \in \Sigma_i$ such that

$$u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}) \text{ for all } s_{-i} \in \Sigma_{-i}$$

Iterated deletion of strictly dominated strategies:

Define a sequence of strategy sets for each player $D_i^0, D_i^1, \ldots$

$D_i^0 = S_i$

$D_i^{k+1} = \{s_i \mid \text{there is no } \sigma_i \in \Sigma_i \text{ such that } u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}) \text{ for all } s_{-i} \in D_{-i}^k\}$

$D_i = lim_{k\to\infty} D_i^k$

As long as G is finite (i.e. n is finite and each S_i is finite), this limit is well-defined and non-empty

Example

	X	*Y*	*Z*
A	3, 2	3, 1	0, 1
B	2, 3	2, 2	3, 1
C	1, 2	1, 2	4, 1
~~*D*~~	~~0, 2~~	~~0, 3~~	~~1, 3~~

→

	X	*Y*	~~*Z*~~
A	3, 2	3, 1	~~0, 1~~
B	2, 3	2, 2	~~3, 1~~
C	1, 2	1, 2	~~4, 1~~

→

	X	Y
A	3, 2	3, 1
B	2, 3	2, 2
~~C~~	~~1, 2~~	~~1, 2~~

→

	X	~~Y~~
A	3, 2	~~3, 1~~
B	2, 3	~~2, 2~~

→

	X
A	3, 2
~~B~~	~~2, 3~~

CKR and IDSDS

An easy result: suppose w is a state in some model of G. If $w \in Rat_i$, then $f(w) \in D_i^1$

If s_i is strictly dominated by some other strategy, σ_i, then whatever beliefs i has about what her opponents will do, σ_i yields higher expected utility than s_i.

Note: even if s_i is strictly dominated only by a mixed strategy for player i, for any given beliefs that i might have, we can find a pure strategy for i which beats s_i.

Proposition 1a: suppose w is a state in some model of G. If $w \in CKR$, then $f(w) \in D$

Intuition: if I know you're rational, then I know you'll play one of the strategies in D_i^1. Thus I'll play one of the strategies in D_j^2…

Example

k levels of knowledge of rationality are required for *k* + 1 rounds of deletion

f_A	*A*	*C*	*C*	*D*
$\mathcal{I}_A$	•	•	•	•
$\mathcal{I}_B$	•	•	•	•
f_B	*Y*	*Y*	*Z*	*Z*
	Rat_1	Rat_1	Rat_1	$\sim Rat_1$
	Rat_2	Rat_2	Rat_2	Rat_2
	$K_1 Rat_2$	$K_1 Rat_2$	$K_1 Rat_2$	$K_1 Rat_2$
	$K_2 Rat_1$	$K_2 Rat_1$	$\sim K_2 Rat_1$	$\sim K_2 Rat_1$
	$K K_2 Rat_1$	$\sim K_1 K_2 Rat_1$	$\sim K_1 K_2 Rat_1$	$\sim K_1 K_2 Rat_1$
	$\sim K_2 K_1 K_2 Rat_1$			

	X	*Y*	*Z*
A	3, 2	3, 1	0, 1
B	2, 3	2, 2	3, 1
C	1, 2	1, 2	4, 1
D	0, 2	0, 3	1, 3

(This example due
to Giacomo Bonanno)

At the first state, we have 2 levels of knowledge of rationality, and 3 rounds of deletion

Rationalizing Strategies

We have just seen that if there is common knowledge in rationality, the players will play strategies which survive IDSDS.

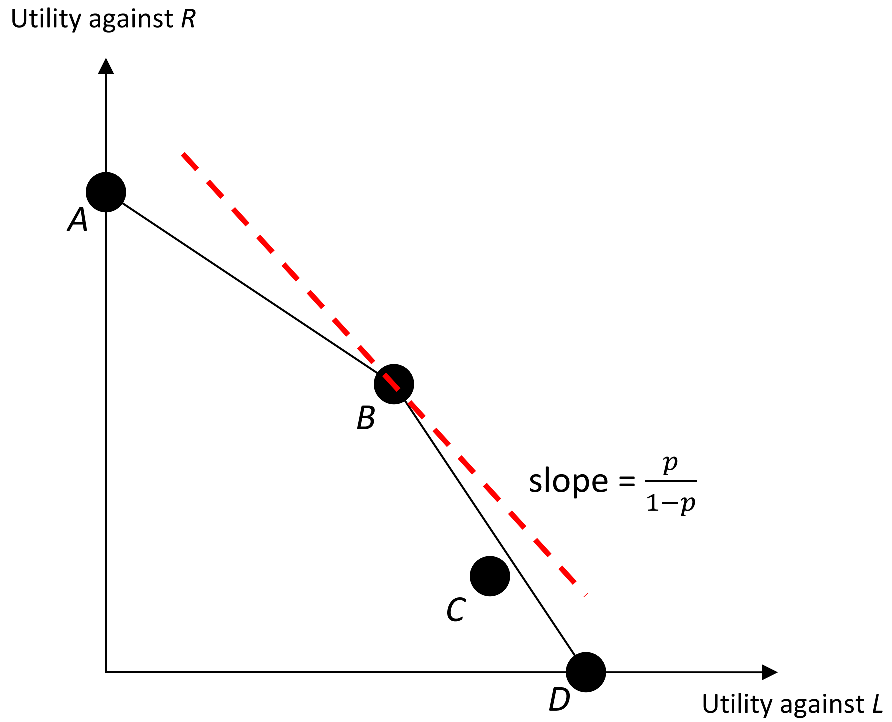
What about the converse: if the players play strategies which survive IDSDS, does mean that there must be common knowledge of rationality?

No! In fact, we've already seen a counterexample (in the Battle of the Sexes game).

However as long as a strategy is not strictly dominated, we can find beliefs which rationalize it, i.e. we can find a state in some model of the game where this strategy is played, and the player is rational. Furthermore:

Proposition 1b: suppose $s \in D$ for game G. Then there is a state w in some model of G such that $w \in CKR$ and $f(w) = s$.

Rationalizing Strategies contd.



	L	R
A	0	5
B	3	3
C	4	1
D	5	0

Player 1's payoffs

Expected utility from choosing strategy s_1

$$p \cdot u_1(s_1, L) + (1 - p) \cdot u_1(s_1, R)$$

(constant along the red dotted line)

Correlated Beliefs

Note: when there are more than two players, a player may need to have correlated beliefs about her opponents' strategies to rationalize a strategy.

Example:

A

	L	*R*
U	1, 1, 3	1, 0, 3
D	0, 1, 0	0, 0, 0

B

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

C

	L	*R*
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

Strategy *B* for player 3 is not strictly dominated. Can it be rationalized?

Try: 3 thinks 1 plays *U* with probability *p* and 2 plays *L* with probability *q* (independent)

$$EU_3(A) = 3p;\ EU_3(B) = 2pq + 2(1-p)(1-q);\ EU_3(C) = 3(1-p)$$

If $p \leq ½$, then *C* beats *B*; if $p \geq ½$, then *A* beats *B*

But if 3 believes $prob(UL) = prob(DR) = \frac{1}{2}$, *B* is clearly optimal

Recap

Proposition 1a: suppose w is a state in some model of G. If $w \in CKR$, then $f(w) \in D$

Proposition 1b: suppose $s \in D$ for game G. Then there is a state w is some model of G such that $w \in CKR$ and $f(w) = s$.

We say that common knowledge of rationality *characterizes* iterated deletion of strictly dominated strategies

Nash equilibrium

A strategy profile $s \in S$ is a (pure strategy) *Nash equilibrium* if, for every player i,

$$u_i(s) \geq u_i(s_i', s_{-i}), \text{for all } s_i' \in S_i.$$

Let N denote the set of Nash equilibrium strategy profiles. It is easy to show that $N \subseteq D$.

Mutual rationality and knowledge in the strategy profile characterize pure strategy Nash equilibrium.

Proposition 2a: suppose w is a state in some model of G. If $w \in Rat \cap K([f(w)])$, then $f(w) \in N$

Proposition 2b: suppose $s \in N$ for some game G. Then there is a state w is some model of G such that $w \in Rat \cap K([f(w)])$ and $f(w) = s$.

Mixed Strategy Nash Equilibrium

Some games that don't have a Nash equilibrium in pure strategies, and even games that do sometimes have mixed strategy equilibria as well.

A strategy profile $\sigma \in \Sigma$ is a (mixed strategy) *Nash equilibrium* if, for every player i,

$$u_i(\sigma) \geq u_i(s_i', \sigma_{-i}), \text{for all } s_i' \in S_i.$$

Matching pennies

	H	*T*
H	1, 0	0, 1
T	0, 1	1, 0

$$\left(\frac{1}{2}H \;\&\; \frac{1}{2}T,\ \frac{1}{2}H \;\&\; \frac{1}{2}T\right)$$

Battle of the sexes

	P	*C*
P	3, 1	0, 0
C	0, 0	1, 3

$$\left(\frac{3}{4}P \;\&\; \frac{1}{4}C,\ \frac{1}{4}P \;\&\; \frac{3}{4}C\right)$$

Modeling Mixed Strategies

In our models of games, each state specifies a unique strategy for each player. How can we model a mixed strategy?

The usual approach is to think of a mixed strategy by player 1 as represented by uncertainty in the mind of player 2 about what player 1 will do.

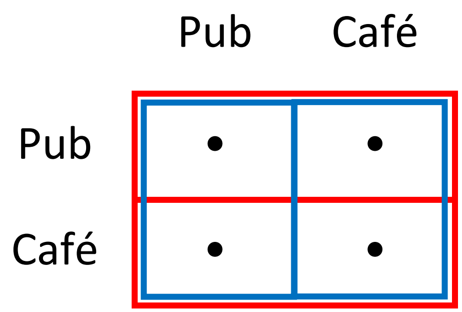
We can think of a mixed strategy Nash equilibrium as an *equilibrium in conjectures*

This interpretation works better in Battle of the Sexes than in Matching Pennies (see Reny & Robson, GEB 2004)

Example

	Pub	Café
Pub	3, 1	0, 0
Café	0, 0	1, 3

G



$\mathcal{J}_1$ and $\mathcal{J}_2$

	Pub	Café
Pub	3/16	9/16
Café	1/16	3/16

$p_1 = p_2 = p$

At every state, player 1 believes that player 2 goes to the Pub with probability ¼ and to the Café with probability ¾

At every state, player 2 believes that player 1 goes to the Pub with probability ¾ and to the Café with probability ¾

Agreeing to Agree

When there are more than two players, we need to be careful. What if player 1 and player 2 have different beliefs about what player 3 will do?

Aumann & Brandenburger (Ecta. 1995) have an elegant way of ruling this out. Suppose:

- the players have a common prior over the state space: $p_1 = \cdots = p_n = p$
- each player's conjecture about her opponents' strategies is common knowledge

Then these conjectures must be the same!

This is an application of Aumann's (1976) agreeing to disagree theorem

Recap

Mixed Strategy Nash equilibrium is characterized by:

2 player case:

- mutual knowledge of rationality
- mutual knowledge of conjectures

3+ player case:

- mutual knowledge of rationality
- common prior
- common knowledge of conjectures

Correlated Equilibrium

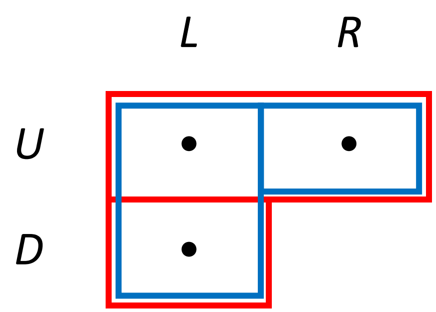
A *correlated equilibrium* is a distribution over (pure) strategy profiles $p \in \Delta S$ such that, for every player i and every s_i with $p(s_i) > 0$

$$\sum_{s_{-i} \in S_{-i}} p(s_{-i} \mid s_i) \cdot u_i(s) \;\geq\; \sum_{s_{-i} \in S_{-i}} p(s_{-i} \mid s_i) \cdot u_i(s_i', s_{-i})\,, \qquad \forall s_i' \in S_i$$

Example:

	L	*R*
U	4, 4	1, 5
D	5, 1	0, 0

G



$\mathcal{J}_1$ and $\mathcal{J}_2$

	L	*R*
U	1/3	1/3
D	1/3	0

p

Characterizing Correlated Equilibrium

Correlated equilibrium is characterized by:

- Common knowledge of rationality
- Common prior

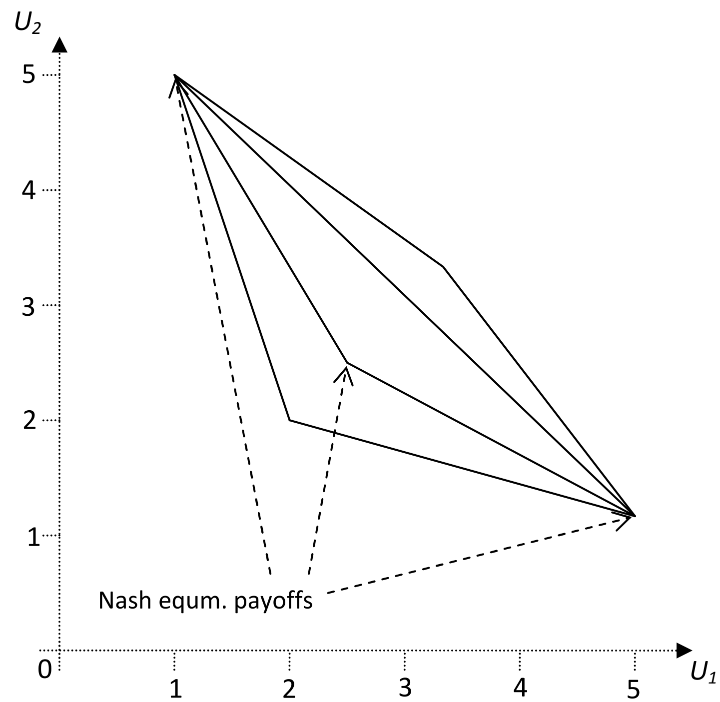
No assumptions about knowledge of strategies or conjectures are needed!

Caveat: the set of correlated equilibria can be very large. We can find:

- Outcomes outside the convex hull of the set of Nash equilibria
- Outcomes where each player does worse than in the unique Nash equilibrium

Example

	L	*R*
U	4, 4	1, 5
D	5, 1	0, 0



Example

	W	*X*	*Y*	*Z*
A	0, 0	2, 0	0, 2	1, 1
B	0, 2	0, 0	2, 0	1, 1
C	2, 0	0, 2	0, 0	1, 1
D	1, 1	1, 1	1, 1	3, 3

Unique Nash equilibrium: (D, Z)

Correlated equilibrium: $p(AX) = p(AY) = p(BW) = p(BY) = p(CW) = p(CX) = \frac{1}{6}$

Common Priors

The characterizations of IDSDS and correlated equilibrium differ only in that the latter requires a common prior

This provides a neat illustration of an implication of the common prior assumption

Example:

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

All four strategy profiles survive IDSDS

Yet, in any correlated equilibrium, we must have $p(UL) = p(DL) = 0$