Reasoning in Games: Players as Programs

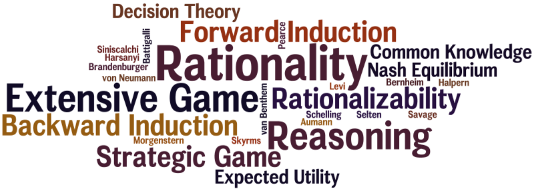
Lecture 4

Eric Pacuit

Department of Philosophy

University of Maryland

pacuit.org



Plan

- ✓ **Monday** Epistemic utility theory, Decision- and game-theoretic background: Nash equilibrium
- ✓ **Tuesday** Introduction to game theory: rationalizability, epistemic game theory, introduction to backward induction
- ✓ **Wednesday** backward and forward induction, Iterated games and learning, Skyrms's model of rational deliberation (introduction);
- **Thursday** Skyrms's model of rational deliberation; *brief* introduction to webppl; Game-theoretic reasoning in webppl; Coordination games (comparing Skyrms's model of deliberation and the webppl approach)
- **Friday** Models of game-theoretic reasoning

Strategic Reasoning

"The word *eductive* will be used to describe a dynamic process by means of which equilibrium is achieved through careful reasoning on the part of the players. Such reasoning will usually require an attempt to simulate the reasoning processes of the other players. Some measure of pre-play communication is therefore implied, although this need not be explicit. To reason along the lines "if I think that he thinks that I think..." requires that information be available on how an opponent thinks."

(pg. 184)

K. Binmore. *Modeling Rational Players*. Economics and Philosophy, 3,179 - 21, 1987.

Deliberational Decision Theory

F. Arntzenius. *No Regrest, or: Edith Piaf Revamps Decision Theory*. Erkenntnis, 68, pgs. 277 - 297, 2008.

J. Joyce. *Regret and Instability in Causal Decision Theory*. Synthese, 187: 1, pgs. 123 - 145, 2012.

I. Douven. *Decision theory and the rationality of further deliberation*. Economics and Philosophy, 18, pgs. 303 - 328, 2002.

Deliberational Decision Theory

Current Evaluation: If Pr_t characterizes your beliefs at time t, then at t you should *evaluate* each act by its (causal, evidential) expected utility computed using Pr_t.

Deliberational Decision Theory

Current Evaluation: If Pr_t characterizes your beliefs at time t, then at t you should *evaluate* each act by its (causal, evidential) expected utility computed using Pr_t.

Full Information: You should act on your time-t utility assessments only if those assessments are based on beliefs that incorporate *all* the evidence that is both freely available to you at t and relevant to the question about what your acts are likely to cause.

Deliberational Decision Theory

Current Evaluation: If Pr_t characterizes your beliefs at time t, then at t you should *evaluate* each act by its (causal, evidential) expected utility computed using Pr_t.

Full Information: You should act on your time-t utility assessments only if those assessments are based on beliefs that incorporate *all* the evidence that is both freely available to you at t and relevant to the question about what your acts are likely to cause.

Sometimes initial opinions fix actions, *but not always* (e.g., Murder Lesion, Psychopath Button)

Deliberation in games

- The Harsanyi-Selten tracing procedure
- Brian Skyrms' model of "dynamic deliberation"
- 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.*

- What operations transform the models?
- Where does the "new information" come from? What are player i's opponents thinking about doing? ("update by emulation")
- Why keep deliberating?

Information Feedback

In the simplest case, deliberation is trivial; one calculates expected utility and maximizes

Information Feedback

In the simplest case, deliberation is trivial; one calculates expected utility and maximizes

Information feedback: “the very process of deliberation may generate information that is relevant to the evaluation of the expected utilities. Then, processing costs permitting, a Bayesian deliberator will feed back that information, modifying his probabilities of states of the world, and recalculate expected utilities in light of the new knowledge.”

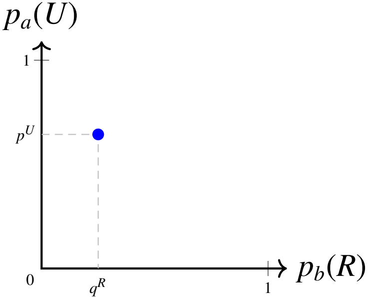
Rational deliberation in games

B. Skyrms (1990). *The Dynamics of Rational Deliberation*. Harvard University Press.

> *It is not just a question of what common knowledge obtains at the moment of truth, but also how common knowledge is preserved, created, or destroyed in the deliberational process which leads up to the moment of truth.* *(pg. 159)*

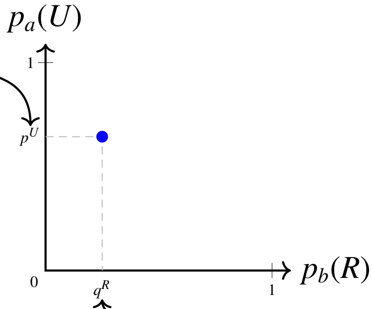
b

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



a's current state of indecision

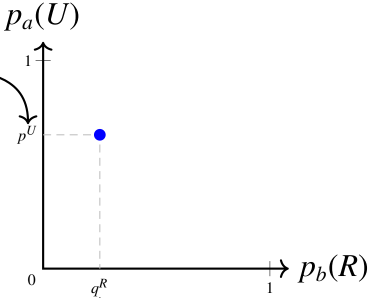
		b	
		L	R
a	U	3,1	0,0
	D	0,0	1,3



a's current belief about what b is going to do

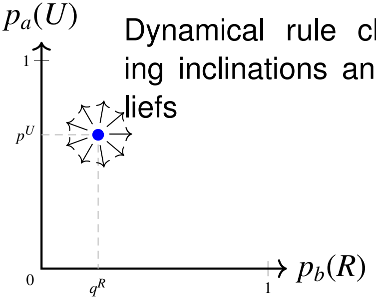
b's current belief about what a is going to do

		b	
		L	R
a	U	3,1	0,0
	D	0,0	1,3



b's current state of indecision

a \ b	L	R
U	3,1	0,0
D	0,0	1,3



Dynamical rule changing inclinations and beliefs

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

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

For each player $i \in \mathcal{A}$, the **state of indecision** is a pair (I_i, P_i), where $I_i \in \Delta(S_i)$ is called i's **inclinations** and $P_i \in \Delta(S_{-i})$ is i's **beliefs** about the other player's choice.

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

For each player $i \in \mathcal{A}$, the **state of indecision** is a pair (I_i, P_i), where $I_i \in \Delta(S_i)$ is called i's **inclinations** and $P_i \in \Delta(S_{-i})$ is i's **beliefs** about the other player's choice.

The **expected utility** of $s \in S_i$ is: $EU_i(s) = \sum_{t \in S_{-i}} P_i(t) u_i(s, t)$.

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

For each player $i \in \mathcal{A}$, the **state of indecision** is a pair (I_i, P_i), where $I_i \in \Delta(S_i)$ is called i's **inclinations** and $P_i \in \Delta(S_{-i})$ is i's **beliefs** about the other player's choice.

The **expected utility** of $s \in S_i$ is: $EU_i(s) = \sum_{t\in S_{-i}} P_i(t)u_i(s,t)$.

The **status quo** is: $SQ_i = \sum_{s_i \in S_i} I_i(s_i)EU_i(s_i)$.

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

$\mathbf{P_a} = \langle 0.2, 0.8 \rangle$ and $\mathbf{P_b} = \langle 0.4, 0.6 \rangle$

$$EU(U) = 0.4 \cdot 2 + 0.6 \cdot 0 = 0.8$$
$$EU(D) = 0.4 \cdot 0 + 0.6 \cdot 1 = 0.6$$
$$EU(L) = 0.2 \cdot 1 + 0.8 \cdot 0 = 0.2$$
$$EU(R) = 0.2 \cdot 0 + 0.8 \cdot 2 = 1.6$$
$$SQ_A = 0.2 \cdot EU(U) + 0.8 \cdot EU(D) = 0.2 \cdot 0.8 + 0.8 \cdot 0.6 = 0.64$$
$$SQ_B = 0.4 \cdot EU(L) + 0.6 \cdot EU(R) = 0.4 \cdot 0.2 + 0.6 \cdot 1.6 = 1.04$$

Nash dynamics

The **covetability** of a strategy s for player i is: $cov_i(s) = \max(EU_i(s) - SQ_i, 0)$.

Nash dynamics

The **covetability** of a strategy s for player i is: $cov_i(s) = \max(EU_i(s) - SQ_i, 0)$.

Then, **Nash dynamics** rule transforms $I_i \in \Delta(S_i)$ into a new probability $I_i' \in \Delta(S_i)$ as follows. For each $s \in S_i$:

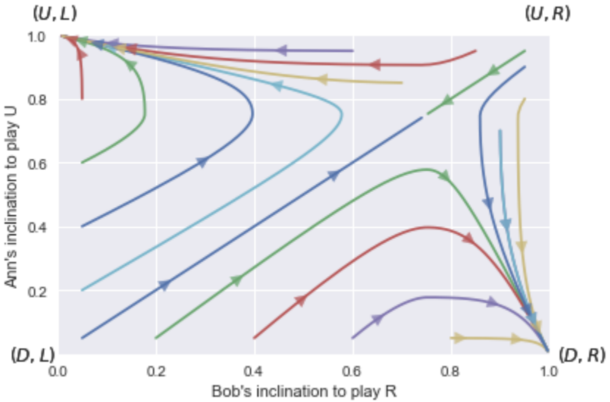
$$I_i'(s) = \frac{k \cdot I_i(s) + cov_i(s)}{k + \sum_{s \in S_i} cov_i(s)},$$

where $k > 0$ is the "index of caution".

Update by emulation

1. The players' initial states of indecision and the dynamical rule used to update inclinations are common knowledge.
2. Each player assumes that the other players are rational deliberators who have just carried out a similar process. So, she can simply go through their calculations to see their new states of indecision and update her beliefs for their acts accordingly.

BoS - Nash Dynamics



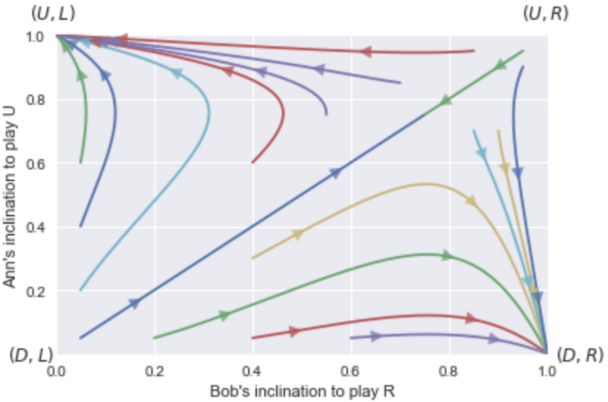
Bayes dynamics

The **Bayes dynamics**, also called **Darwin dynamics**, transforms $I_i \in \Delta(S_i)$ into a new probability $I_i' \in \Delta(S_i)$ as follows. For each $s \in S_i$:

$$I_i'(s) = I_i(s) + \frac{1}{k} I_i(s) \frac{EU_i(s) - SQ_i}{SQ_i}.$$

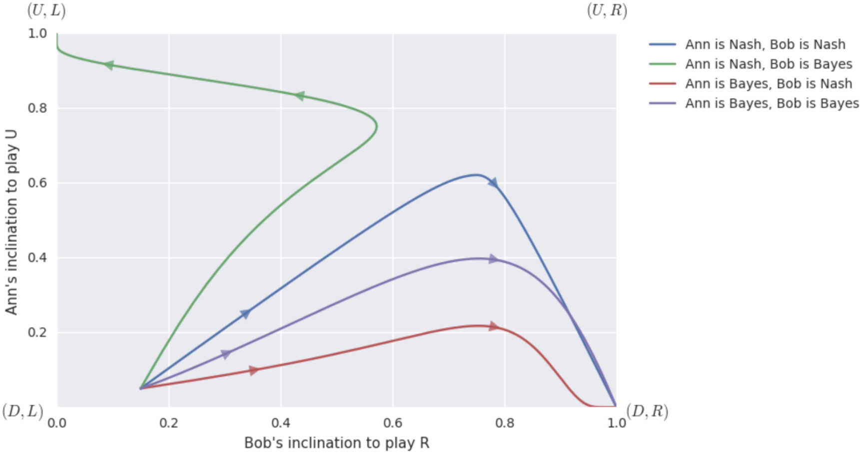
where $k > 0$ is the "index of caution".

BoS - Bayes

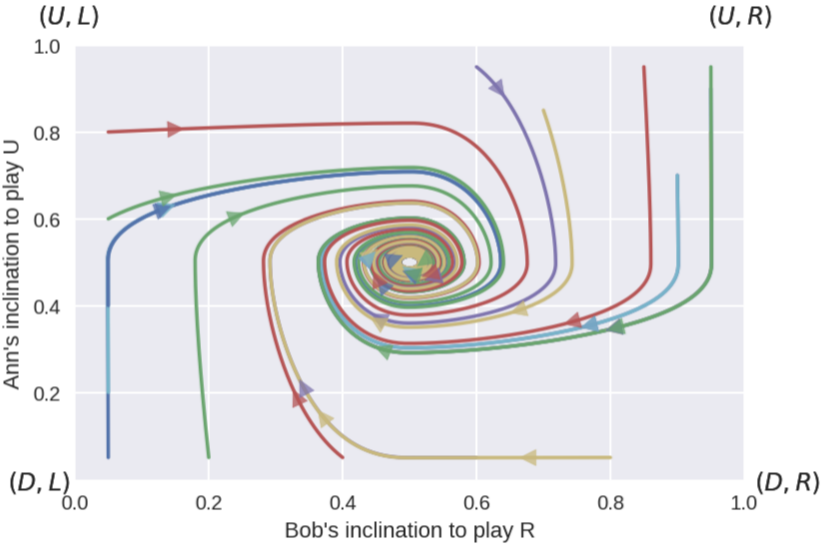


(U, L)
(U, R)
(D, L)
(D, R)
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
Ann's inclination to play U
Bob's inclination to play R

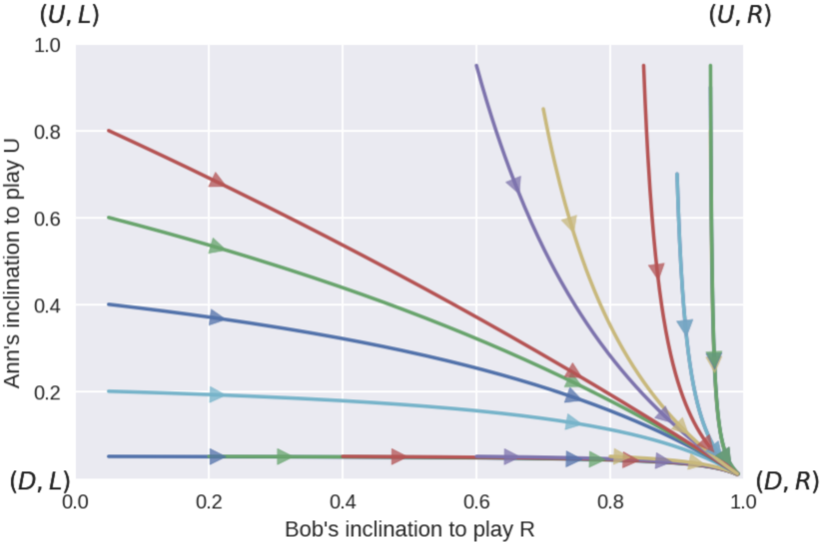
Battle of the sexes



Matching pennies - Nash deliberators



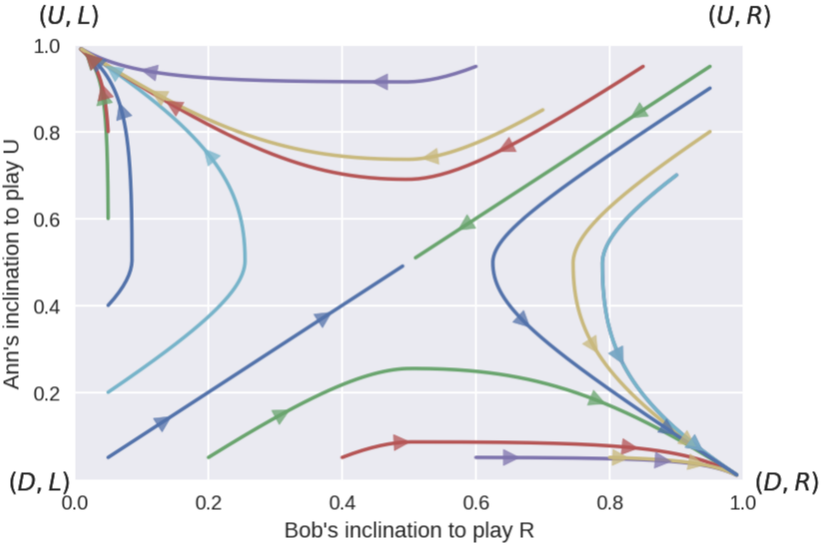
Prisoner's dilemma - Nash deliberators



Stag hunt

		Bob	
		S	H
Ann	S	3, 3	0, 2
	H	2, 0	1, 1

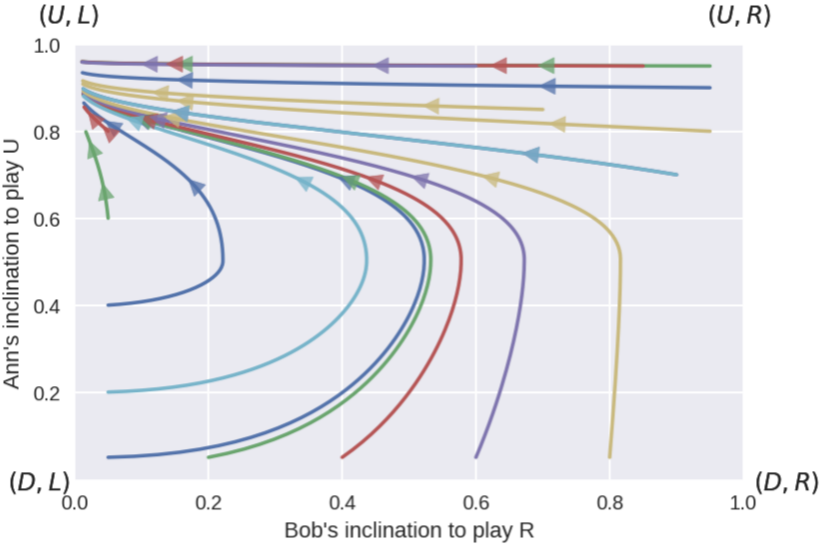
Stag hunt - Nash deliberators



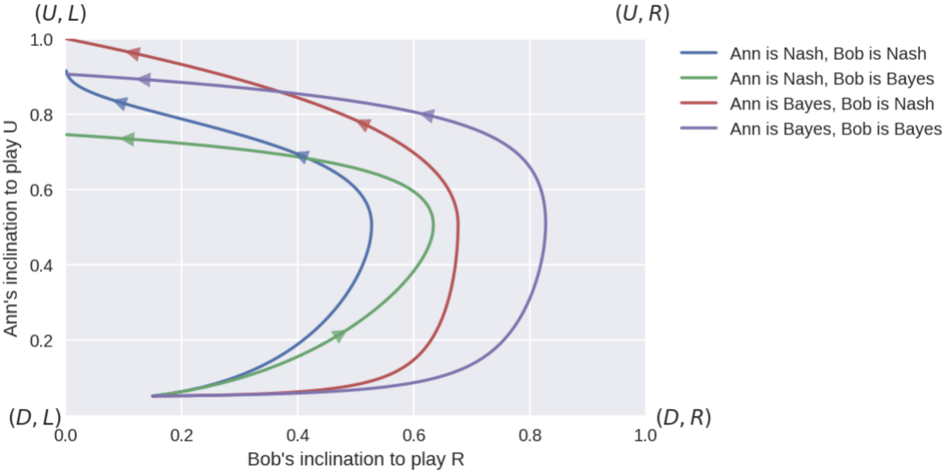
Samuelson's game

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

Samuelson game - Nash deliberators



Samuelson game



Learning to Play

Theorem. If players start with subjectively rational strategies, and if their individual subjective beliefs regarding opponents' strategies are "compatible with truly chosen strategies", then they must converge in a finite amount of time to play according to an ϵ-Nash in the repeated game.

E. Kalai and E. Lehrer. *Rational Learning Leads to Nash Equilibrium*. Econometrica, 61:5, pgs. 1019 - 1045, 1993.

Y. Shoham, R. Powers and T. Granager. *If multi-agent learning is the answer, what is the question?*. Artificial Intelligence, 171(7), pgs. 365 - 377, 2007.

Modeling Deliberation in Games

- Characterize outcomes in terms of *accessibility* and/or *stability*
- Relation with *correlated equilibrium* (correlation through rational deliberation)

Modeling Deliberation in Games

- Characterize outcomes in terms of *accessibility* and/or *stability*
- Relation with *correlated equilibrium* (correlation through rational deliberation)
- Comparison with other models of deliberation in games (categorize pure strategies)

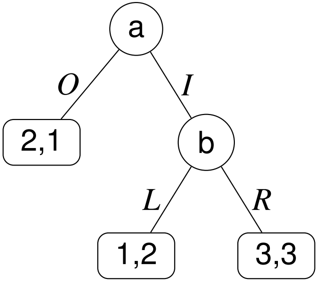
Modeling Deliberation in Games

- Characterize outcomes in terms of *accessibility* and/or *stability*
- Relation with *correlated equilibrium* (correlation through rational deliberation)
- Comparison with other models of deliberation in games (categorize pure strategies)
- Generalize the basic model: extensive games (with imperfect information), imprecise probabilities, more than two players
- Weaken the common knowledge assumptions (payoffs, beliefs, dynamical rule, updating by emulation)

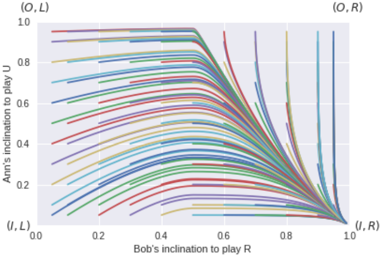
Modeling Deliberation in Games

- Characterize outcomes in terms of *accessibility* and/or *stability*
- Relation with *correlated equilibrium* (correlation through rational deliberation)
- Comparison with other models of deliberation in games (categorize pure strategies)
- Generalize the basic model: extensive games (with imperfect information), imprecise probabilities, more than two players
- Weaken the common knowledge assumptions (payoffs, beliefs, dynamical rule, updating by emulation)
- Deliberation in decision theory ("deliberation crowds out prediction", logical omniscience)

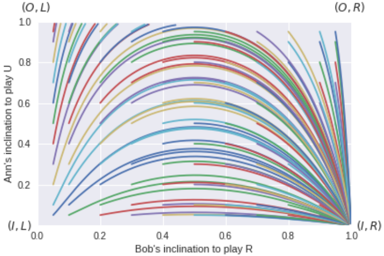
Deliberation on extensive games



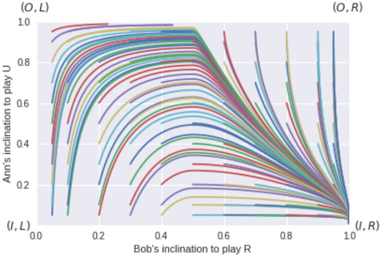
	L	R
O	$2, 1$	$2, 1$
I	$1, 2$	$3, 3$



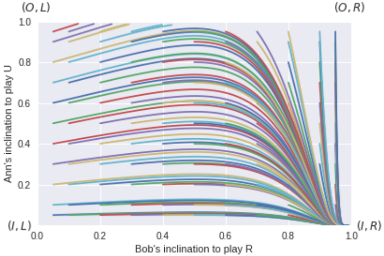
Ann is Nash, Bob is Nash



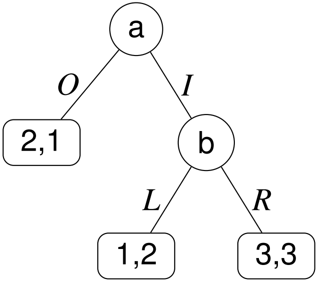
Ann is Bayes, Bob is Bayes



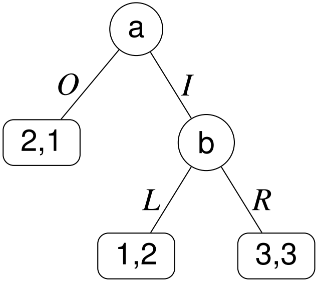
Ann is Nash, Bob is Bayes



Ann is Bayes, Bob is Nash

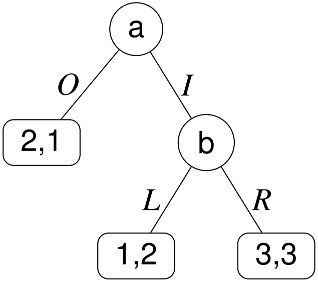


	L	R
O	2, 1	2, 1
I	1, 2	3, 3



	L	R
O	$2, 1$	$2, 1$
I	$1, 2$	$3, 3$

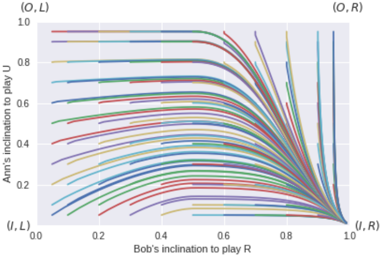
When the players deliberate simultaneously, Bob's expected utility of L is a weighted average of his payoff if Ann chooses O and his payoff if Ann chooses I.



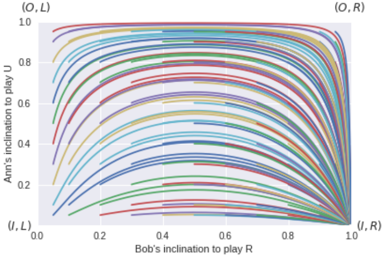
	L	R
O	$2, 1$	$2, 1$
I	$1, 2$	$3, 3$

When the players deliberate simultaneously, Bob's expected utility of L is a weighted average of his payoff if Ann chooses O and his payoff if Ann chooses I.

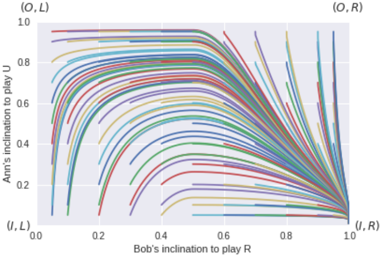
However, when deliberating on the extensive form game, Bob should calculate the expected utilities by conditioning on his information at his decision node: Bob should assign probability 0 to Ann choosing O, and this does not change during deliberation.



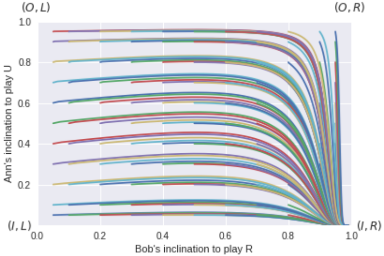
Ann is Nash, Bob is Nash



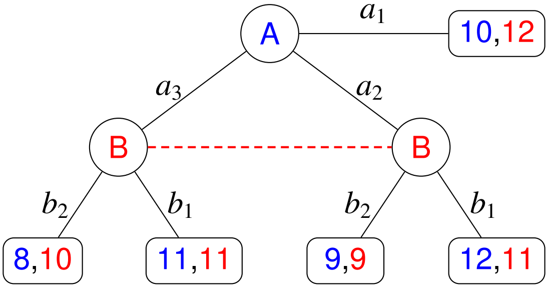
Ann is Bayes, Bob is Bayes



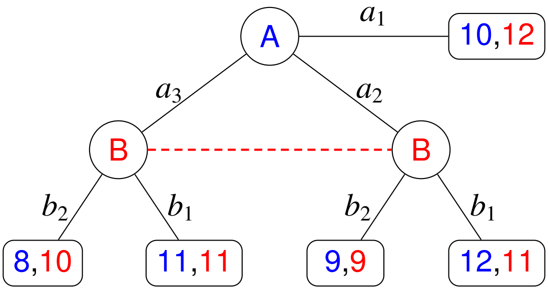
Ann is Nash, Bob is Bayes



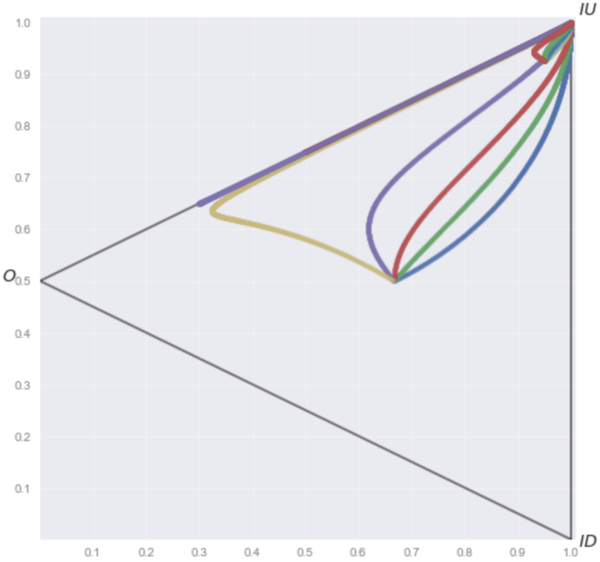
Ann is Bayes, Bob is Nash

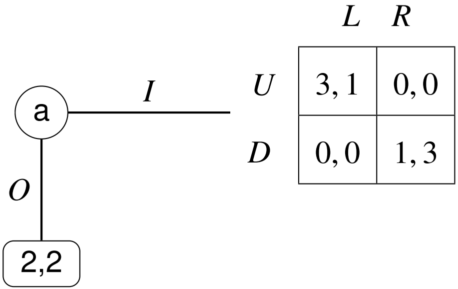


	b_1 if a_2 or a_3	b_2 if a_2 or a_3
a_1	10,12	10,12
a_2	12,11	9,9
a_3	11,11	8,10

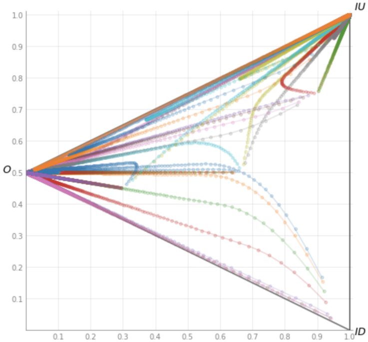


- No matter what Ann's probabilities are for playing a_2 and a_3, Bob is always better off playing b_1.
- Thus, Bob will play b_1 at his information set
- Knowing this, Ann will play a_2
- Dynamic deliberation will never lead to the "bad" equilibrium (a_1, b_2 if a_2 or a_3)

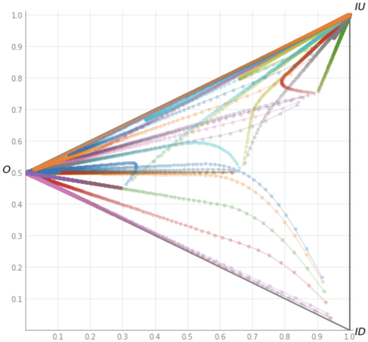

IU
O
ID
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0



	L	R
O	2, 2	2, 2
IU	3, 1	0, 0
ID	0, 0	1, 3



Nash deliberators



Bayes deliberators

Note that both Bayes and Nash deliberators converge on (IU, L) and (O, ID).

Note that both Bayes and Nash deliberators converge on (IU, L) and (O, ID).

If Bob is a backward induction reasoner, then he ignores Ann's initial move as he deliberates between L and R.

Note that both Bayes and Nash deliberators converge on (IU, L) and (O, ID).

If Bob is a backward induction reasoner, then he ignores Ann's initial move as he deliberates between L and R.

On the other hand, if Bob is a forward induction reasoner, then, during deliberation, he should assign probability 0 to Ann choosing I then D (since it is strictly dominated by choosing O). This belief about Ann's choice does not change during deliberation.

J. McKenzie Alexander. *Local interactions and the dynamics of rational deliberation*. Philosophical Studies 147 (1), 2010.

Consider a social network $\langle N, E \rangle$ (connected graph)

Consider a social network $\langle N, E \rangle$ (connected graph)

Convention: If there is a directed edge from A to B, then A always plays row and B always play column, and the interactions of Row and Column are symmetric in the available strategies.

Consider a social network $\langle N, E \rangle$ (connected graph)

Convention: If there is a directed edge from A to B, then A always plays row and B always play column, and the interactions of Row and Column are symmetric in the available strategies.

Let $\nu_i = \{i_1, \ldots i_j\}$ be i's neighbors

Consider a social network $\langle N, E \rangle$ (connected graph)

Convention: If there is a directed edge from A to B, then A always plays row and B always play column, and the interactions of Row and Column are symmetric in the available strategies.

Let $\nu_i = \{i_1, \ldots i_j\}$ be i's neighbors

$\mathbf{p'_{a,b}(t+1)}$ is represents the incremental refinement of player a's state of indecision given his knowledge about player b's state of indecision (at time $t+1$).

Consider a social network $\langle N, E \rangle$ (connected graph)

Convention: If there is a directed edge from A to B, then A always plays row and B always play column, and the interactions of Row and Column are symmetric in the available strategies.

Let $\nu_i = \{i_1, \ldots i_j\}$ be i's neighbors

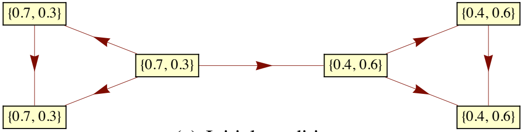
$\mathbf{p'_{a,b}(t+1)}$ is represents the incremental refinement of player a's state of indecision given his knowledge about player b's state of indecision (at time $t+1$).

Pool this information to form your new probabilities:

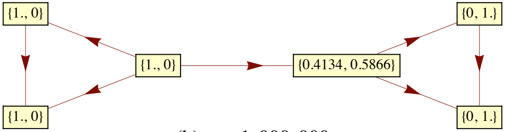
$$\mathbf{p}_i(t+1) = \sum_{j=1}^{k} w_{i,i_j} \mathbf{p}'_{i,i_j}(t+1)$$

		Billy	
		Boxing	Ballet
Maggie	Boxing	$(2,1)$	$(0,0)$
	Ballet	$(0,0)$	$(1,2)$

Fig. 7 The game of Battle of the Sexes.



(a) Initial conditions



(b) $t = 1,000,000$

Fig. 8 Battle of the Sexes played by Nash deliberators ($k = 25$) on two cycles connected by a bridge edge (values rounded to the nearest 10^{-4}).

- Allowing for local interactions in the dynamics of rational deliberation breaks the link between convergent points of the deliberative dynamics and Nash equilibrium points of the underlying game.
- It is no longer true that all dynamical rules have fixed points that maximize expected utility of the status quo.
- The effect of local interactions reveals reasons for preferring the Bayesian dynamics over the Nash dynamics.

Deliberation in games

- The Harsanyi-Selten tracing procedure
- Brian Skyrms' model of "dynamic deliberation"
- Robin Cubitt and Robert Sugden's "reasoning based expected utility procedure"
- Johan van Benthem et col.'s "virtual rationality *announcements*"

EP. *Dynamic models of rational deliberation in games*. in *Strategic Reasoning*, van Benthem, Gosh, and Verbrugge, ed., 2015.

Introduction to webppl for coordination games