Computational Game Theory in Julia

Eric Pacuit, University of Maryland

Lecture 4

ESSLLI 2023

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.*

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

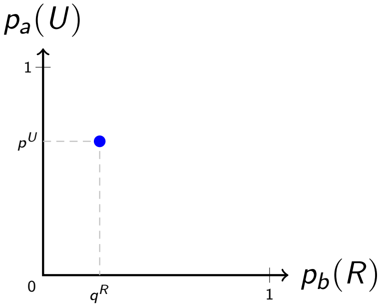
Information feedback: "the very process of deliberation may generate information that is relevant to the evaluation of the expected utilities. Then, processing costs permitting, a Bayesian deliberator will feed back that information, modifying his probabilities of states of the world, and recalculate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

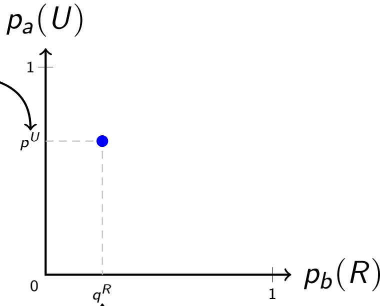
		b	
		L	R
a	U	3,1	0,0
	D	0,0	1,3



a's current state of indecision

b

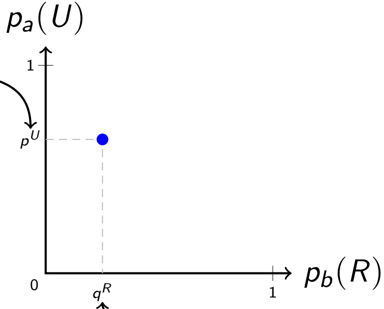
a		L	R
	U	3,1	0,0
	D	0,0	1,3



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

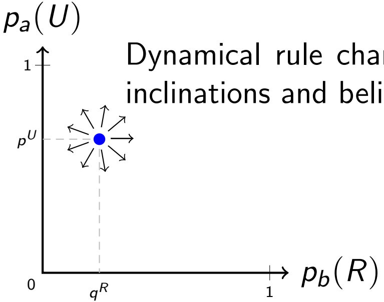
b's current belief about what a is going to do

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



b's current state of indecision

		b	
		L	R
a	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 $i \in N$, 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 $i \in N$, 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 $i \in N$, 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

Ann		L	R
	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 $s \in S_i$:

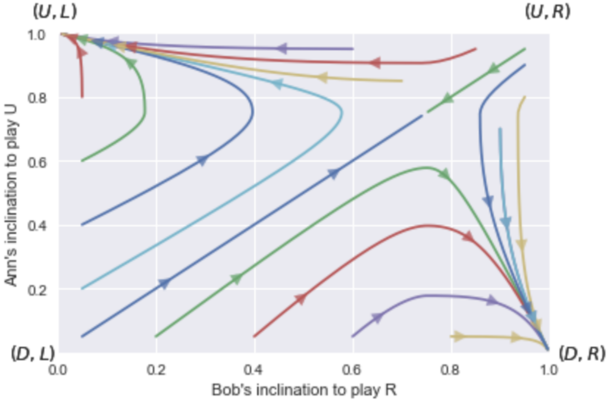
$$I_i'(s) = \frac{k \cdot I_i(s) + cov_i(s)}{k + \sum_{s \in S_i} cov_i(s)},$$

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



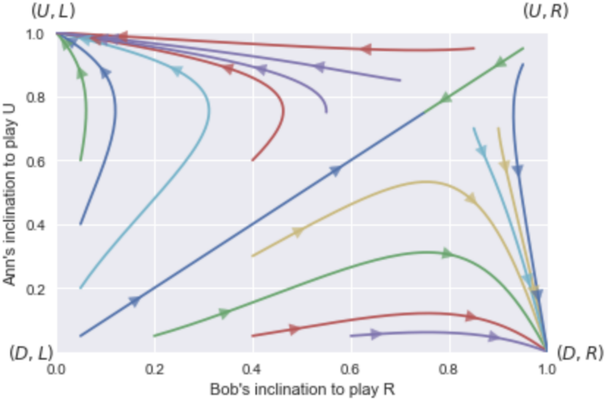
Bayes dynamics

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 $s \in S_i$:

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

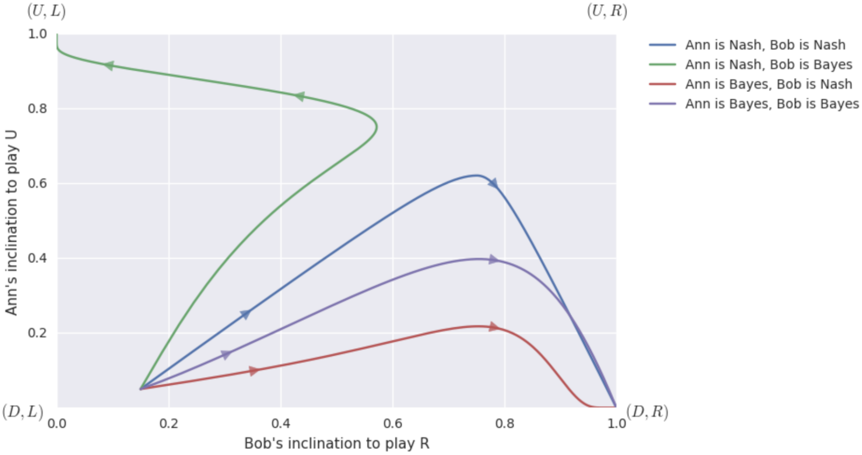
where $k > 0$ is the "index of caution".

BoS - Bayes

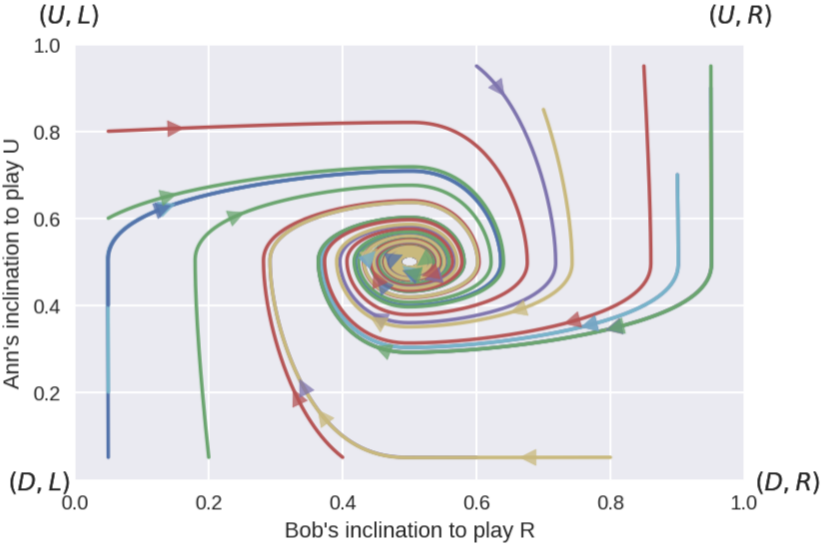


(U, L)
(U, R)
(D, L)
(D, R)
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
Ann's inclination to play U
Bob's inclination to play R

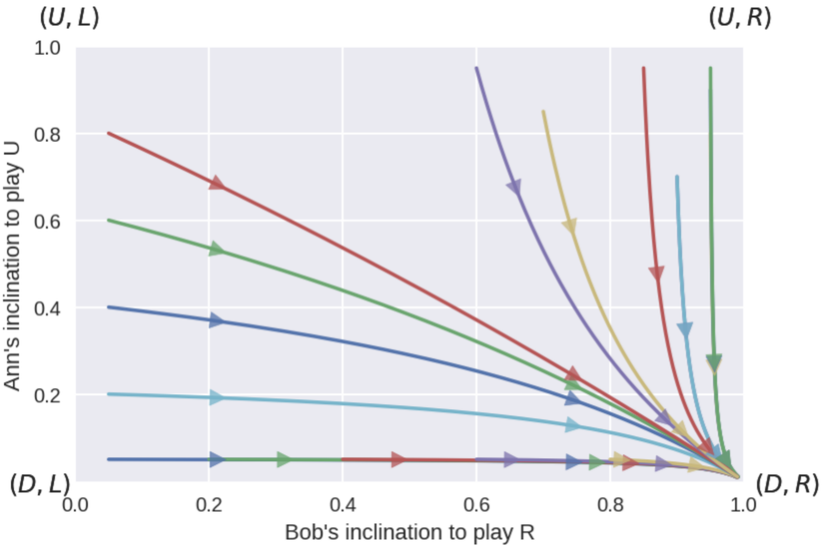
Battle of the sexes



Matching pennies - Nash deliberators



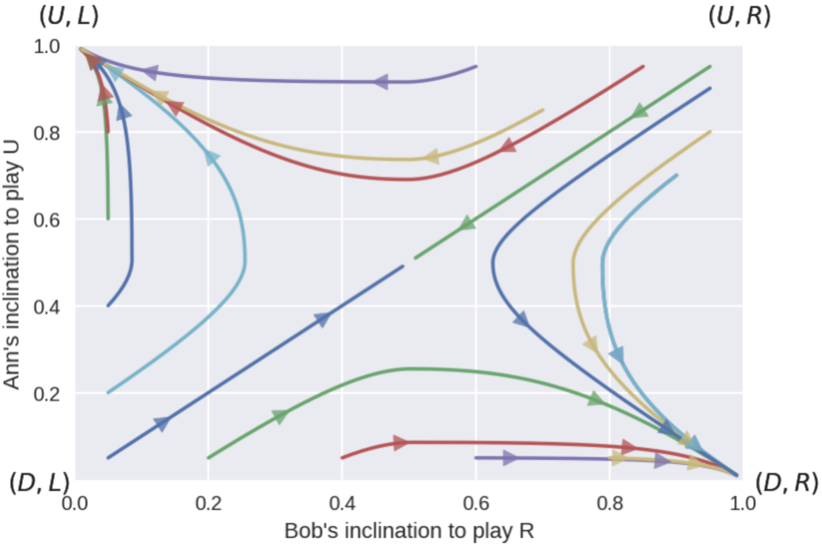
Prisoner's dilemma - Nash deliberators



Stag hunt

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

Stag hunt - Nash deliberators



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*
- Deliberation in decision theory ("deliberation crowds out prediction", logical omniscience)
- Weaken the common knowledge assumptions (payoffs, beliefs, dynamical rule, updating by emulation)

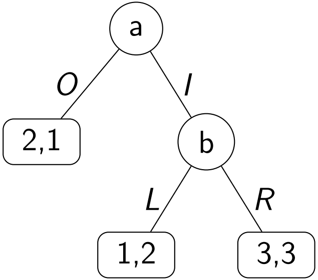
Modeling Deliberation in Games

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

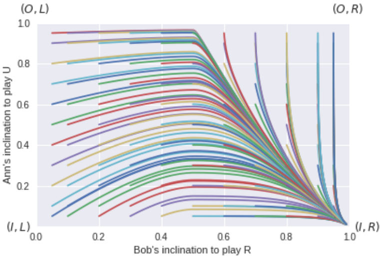
Modeling Deliberation in Games

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

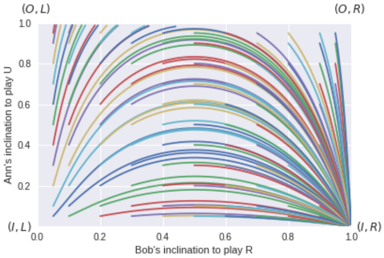
Deliberation on extensive games



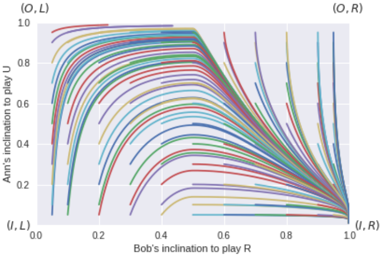
	L	R
O	2, 1	2, 1
I	1, 2	3, 3



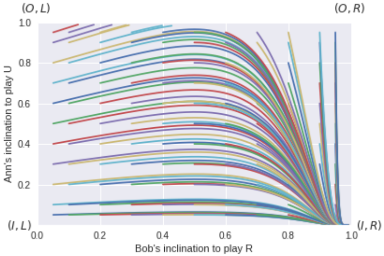
Ann is Nash, Bob is Nash



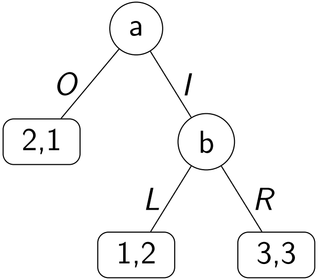
Ann is Bayes, Bob is Bayes



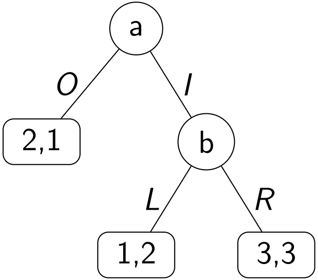
Ann is Nash, Bob is Bayes



Ann is Bayes, Bob is Nash

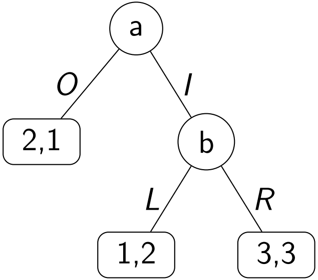


	L	R
O	2, 1	2, 1
I	1, 2	3, 3



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

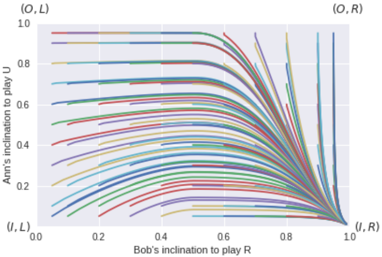
When the players deliberate simultaneously, Bob's expected utility of L is a weighted average of his payoff if Ann chooses O and his payoff if Ann chooses I.



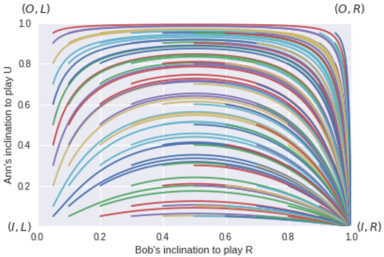
	L	R
O	2, 1	2, 1
I	1, 2	3, 3

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.

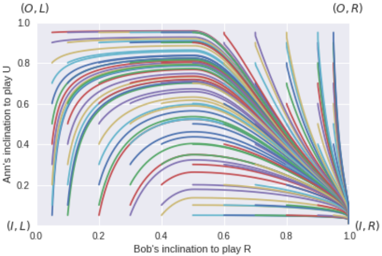
However, when deliberating on the extensive form game, Bob should calculate the expected utilities by conditioning on his information at his decision node: Bob should assign probability 0 to Ann choosing O, and this does not change during deliberation.



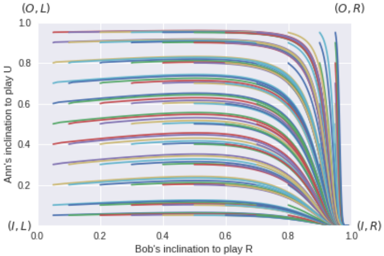
Ann is Nash, Bob is Nash



Ann is Bayes, Bob is Bayes



Ann is Nash, Bob is Bayes



Ann is Bayes, Bob is Nash

Deliberational Dynamics

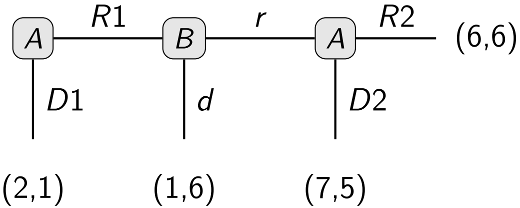
"There is nothing in the nature of deliberational dynamics that requires that deliberators be simpleminded, but the illustrations I have chosen...are relatively unsophisticated. These players follow their noses in the direction of the current apparent good, with no real memory of where they have been, no capability of recognizing patterns, and no sense of where they are going." (Skyrms, pg. 152)

Backward and forward induction reasoning

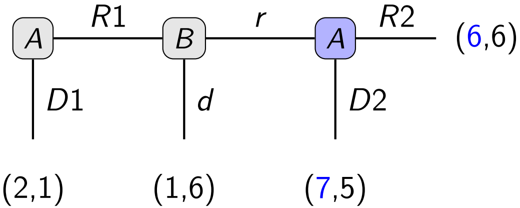
Backward induction reasoning: player's ignore past behavior and reason only about their opponents' future moves.

Forward induction reasoning: player's rationalize past behavior and use it as a basis to form beliefs about their opponents' future moves.

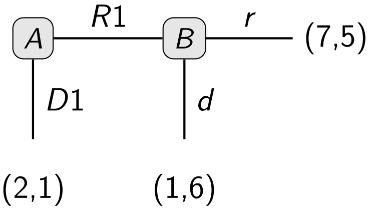
BI Puzzle



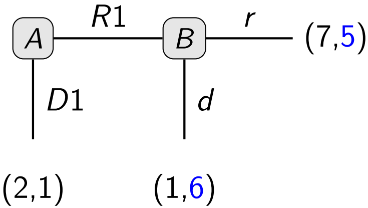
BI Puzzle



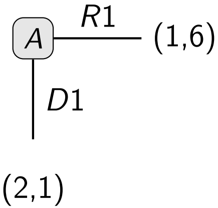
BI Puzzle



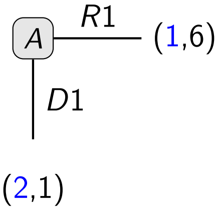
BI Puzzle



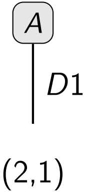
BI Puzzle



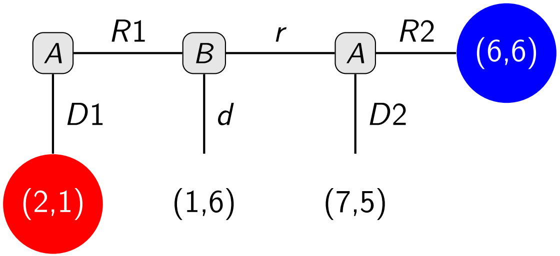
BI Puzzle



BI Puzzle



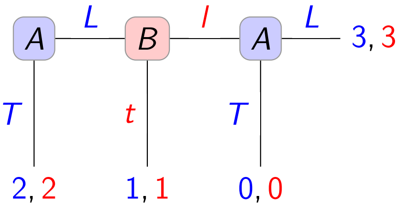
BI Puzzle

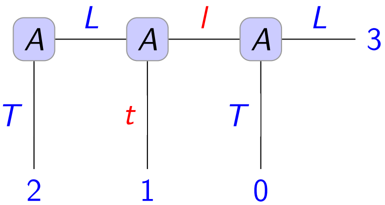


R. Aumann. *Backwards induction and common knowledge of rationality*. Games and Economic Behavior, 8, pp. 6 - 19, 1995.

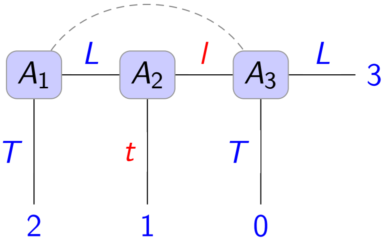
R. Stalnaker. *Knowledge, belief and counterfactual reasoning in games*. Economics and Philosophy, 12, pp. 133 - 163, 1996.

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

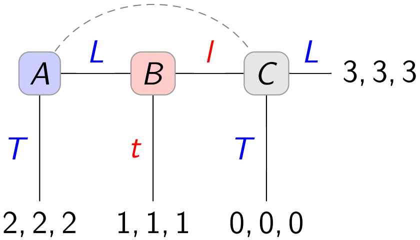

A
L
B
l
A
L
3, 3
T
t
T
2, 2
1, 1
0, 0

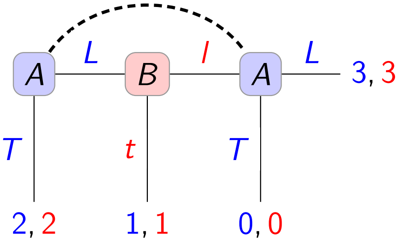


A
A
A
L
l
L
3
T
t
T
2
1
0



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


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

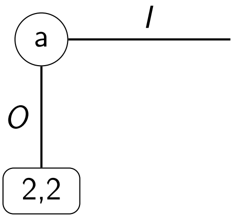


A
L
B
l
A
L
3, 3
T
t
T
2, 2
1, 1
0, 0

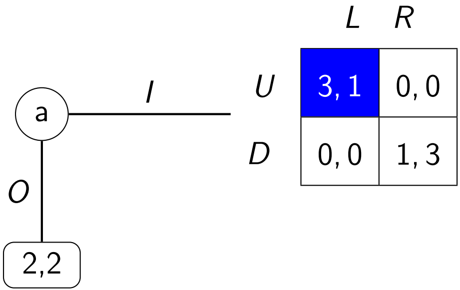
Forward Induction Principle: a player should use all information she acquired about her opponents' past behavior in order to improve her prediction of their future simultaneous and past (unobserved) behavior, relying on the assumption that they are rational.

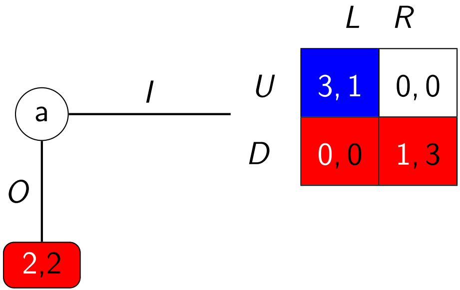
P. Battigalli. *On Rationalizability in Extensive Games*. Journal of Economic Theory, 74, pgs. 40 - 61, 1997.

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

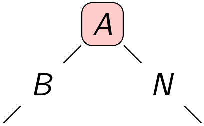


	L	R
U	3, 1	0, 0
D	0, 0	1, 3


L
R
U
3, 1
0, 0
D
0, 0
1, 3
a
I
O
2,2


a
I
O
2,2
L R
U 3, 1 0, 0
D 0, 0 1, 3

Burning Money



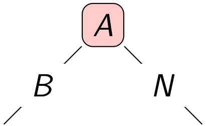
Bob

Ann		l	r
	u	2, 1	-1, 0
	d	-1, 0	0, 3

Bob

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

Burning Money



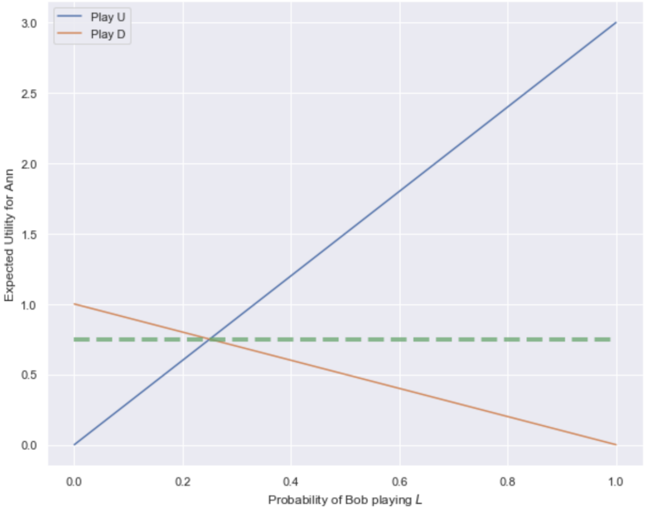
Bob

Ann	l	r
u	2, 1	-1, 0
d	-1, 0	0, 3

Bob

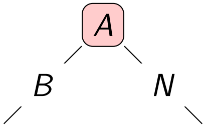
Ann	l	r
u	3, 1	0, 0
d	0, 0	1, 3

1. Ann can achieve at least 0.75 by choosing *N*.
2.
3.
4.
5.
6.



Whatever her belief about Bob, Ann can achieve at least 0.75 by choosing N.

Burning Money



Bob (after *B*)

		l	r
Ann	u	2, 1	-1, 0
	d	-1, 0	0, 3

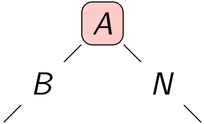
Bob (after *N*)

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

1. Ann can achieve at least 0.75 by choosing *N*.
2. If Ann plays *B*, then she will not subsequently choose *d*.
3. Since Bob knows 2., if Ann plays *B*, Bob will play *l*.
4.
5.

6.

Burning Money



Bob (after B)

Ann		l	r
	u	2, 1	-1, 0
	d	-1, 0	0, 3

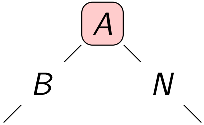
Bob (after N)

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

1. Ann can achieve at least 0.75 by choosing N.
2. If Ann plays B, then she will not subsequently choose d.
3. Since Bob knows 2., if Ann plays B, Bob will play l.
4. So, Ann expects a payout of 2 by playing B.
5.

6.

Burning Money



Bob

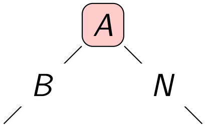
Ann		l	r
	u	2, 1	-1, 0
	d	-1, 0	0, 3

Bob

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

1. Ann can achieve at least 0.75 by choosing *N*.
2. If Ann plays *B*, then she will not subsequently choose *d*.
3. Since Bob knows 2., if Ann plays *B*, Bob will play *l*.
4. So, Ann expects a payout of 2 by playing *B*.
5. If Bob observes *N*, then he knows Ann will not play *d* (since *B* strictly dominances *Nd*).
6.

Burning Money



Left branch (B): Bob / Ann

Ann	l	r
u	2, 1	-1, 0
d	-1, 0	0, 3

Right branch (N): Bob / Ann

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

1. Ann can achieve at least 0.75 by choosing *N*.
2. If Ann plays *B*, then she will not subsequently choose *d*.
3. Since Bob knows 2., if Ann plays *B*, Bob will play *l*.
4. So, Ann expects a payout of 2 by playing *B*.
5. If Bob observes *N*, then he knows Ann will not play *d* (since *B* strictly dominances *Nd*).
6. Knowing all of the above, Ann will play *Nu* and Bob will play *ll*.

Ben-Porath and Dekel (1992) generalized this result as follows: in games in which a player has a strict preference for a equilibrium point, and if this player can self-sacrifice (burning utility), then, based on the forward induction rationality and iterative elimination of weakly dominated strategies, such player will achieve her most preferred outcome.

E. Ben-Porath and E. Dekel. *Signaling future actions and the potential for sacrifice*. Journal of Economic Theory, 57, 36-51, 1992.

P. Battigalli and M. Siniscalchi. *Strong Belief and Forward Induction Reasoning*. Journal of Economic Theory, 106, pp. 356 - 391, 2002.