Models of Strategic Reasoning
Lecture 2

Eric Pacuit

University of Maryland, College Park
ai.stanford.edu/~epacuit

August 7, 2012

Lecture 1: Introduction, Motivation and Background

Lecture 2: The Dynamics of Rational Deliberation

Lecture 3: Reasoning to a Solution: Common Modes of Reasoning in Games

Lecture 4: Reasoning to a Model: Iterated Belief Change as Deliberation

Lecture 5: Reasoning in Specific Games: Experimental Results

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

Suppose that one deliberates by calculating expected utility.

Suppose that one deliberates by calculating expected utility.

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

Suppose that one deliberates by calculating expected utility.

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."

Deliberational Equilibrium

The decision maker cannot decide to do an act that is not an equilibrium of the deliberational process.

(*provided we neglect processing costs...the implementations use a "satisficing level"*)

Deliberational Equilibrium

The decision maker cannot decide to do an act that is not an equilibrium of the deliberational process.

(*provided we neglect processing costs...the implementations use a "satisficing level"*)

This sort of equilibirium requirement can be seen *as a consequence of the expected utility principle* (dynamic coherence).

It is usually neglected because the process of informational feedback is usually neglected.

A Bayesian has to choose between n acts: s_1, s_2, ..., s_n

A Bayesian has to choose between n acts: s_1, s_2, ..., s_n

state of indecision: $\mathbf{P} = \langle p_1, \ldots, p_n \rangle$ of probabilities for each act ($\sum_i p_i = 1$). The *default mixed act* is the mixed act corresponding to the state of indecision (decision makers always make a decision).

A Bayesian has to choose between n acts: s_1, s_2, ..., s_n

state of indecision: $\mathbf{P} = \langle p_1, \ldots, p_n \rangle$ of probabilities for each act ($\sum_i p_i = 1$). The *default mixed act* is the mixed act corresponding to the state of indecision (decision makers always make a decision).

status quo: $EU(\mathbf{P}) = \sum_i p_i \cdot u_i(s_i)$

A person's state of indecision evolves during deliberation.

After computing expected utility, she will believe more strongly that she will ultimately do the acts (or one of those acts) that are ranked more highly than her current state of indecision.

A person's state of indecision evolves during deliberation.
After computing expected utility, she will believe more strongly that she will ultimately do the acts (or one of those acts) that are ranked more highly than her current state of indecision.

Why not just do the act with highest expected utility?

A person's state of indecision evolves during deliberation.
After computing expected utility, she will believe more strongly that she will ultimately do the acts (or one of those acts) that are ranked more highly than her current state of indecision.

Why not just do the act with highest expected utility? *On pain of incoherence*, the player will continue to deliberate if she believes that she is in an informational feedback situation and if she assigns any positive probability at all to the possibility that informational feedback may lead her ultimately to a different decision.

A person's state of indecision evolves during deliberation.
After computing expected utility, she will believe more strongly that she will ultimately do the acts (or one of those acts) that are ranked more highly than her current state of indecision.

Why not just do the act with highest expected utility? *On pain of incoherence*, the player will continue to deliberate if she believes that she is in an informational feedback situation and if she assigns any positive probability at all to the possibility that informational feedback may lead her ultimately to a different decision.

The decision maker follows a "simple dynamical rule" for "making up one's mind"

Seeks the good

The dynamical rule *seeks the good*:

1. the rule raises the probability of an act only if that act has utility greater than the status quo
2. the rule raises the sum of the probability of all acts with utility greater than the status quo (if any)

Seeks the good

The dynamical rule *seeks the good*:

1. the rule raises the probability of an act only if that act has utility greater than the status quo
2. the rule raises the sum of the probability of all acts with utility greater than the status quo (if any)

all dynamical rules that seek the good have the same fixed points: those states in which the expected utility of the status quo is maximal.

Nash Dynamics

covetability of act A: given a state of indecision $\mathbf{P}$
$cov(A) = \max(EU(A) - EU(\mathbf{P}), 0)$

Nash Dynamics

covetability of act A: given a state of indecision $\mathbf{P}$
$cov(A) = \max(EU(A) - EU(\mathbf{P}), 0)$

Nash map: $\mathbf{P} \mapsto \mathbf{P}'$ where each component p_i' is calculated as follows:

$$p_i' = \frac{p_i + cov(A_i)}{1 + \sum_i cov(A_i)}$$

Nash Dynamics

covetability of act A: given a state of indecision $\mathbf{P}$
$cov(A) = \max(EU(A) - EU(\mathbf{P}), 0)$

Nash map: $\mathbf{P} \mapsto \mathbf{P}'$ where each component p_i' is calculated as follows:

$$p_i' = \frac{p_i + cov(A_i)}{1 + \sum_i cov(A_i)}$$

More generally, for $k > 0$,

$$p_i' = \frac{k \cdot p_i + cov(A_i)}{k + \sum_i cov(A_i)}$$

where k is the "index of caution". The higher the k the more slowly the decision maker moves in the direction of acts that look more attractive than the status quo.

decision maker's personal state: $\langle x, y\rangle$ where x is the state of indecision and the probabilities she assigns to the "states of nature" *decision maker's personal state*: $\langle x, y \rangle$ where x is the state of indecision and the probabilities she assigns to the "states of nature"

Dynamics: $\varphi(\langle x, y \rangle) = \langle x', y' \rangle$ consisting of

1. An "adaptive dynamic map" D sending $\langle x, y \rangle$ to x'
2. the informational feedback process I sending $\langle x, y \rangle$ to y'

decision maker's personal state: $\langle x, y \rangle$ where x is the state of indecision and the probabilities she assigns to the "states of nature"

Dynamics: $\varphi(\langle x, y \rangle) = \langle x', y' \rangle$ consisting of

1. An "adaptive dynamic map" D sending $\langle x, y \rangle$ to x'
2. the informational feedback process I sending $\langle x, y \rangle$ to y'

A personal state $\langle x, y \rangle$ is a **deliberational equilibrium** iff $\varphi(\langle x, y \rangle) = \langle x, y \rangle$

Fact. If D seeks the good and I is continuous, then there is a delbierational equilibrium, $\langle x, y\rangle$, for $\langle D, I\rangle$. If D' also seeks the good, then $\langle x, y\rangle$ is also a deliberational equilibrium for $\langle D', I\rangle$. The default mixed act corresponding to x maximizes expected utility at $\langle x, y\rangle$.

Games played by Bayesian deliberators

For each player, the decisions of the other players constitute the relevant state of the world, which together with her decision, determines the consequences in accordance with the payoff matrix.

Games played by Bayesian deliberators

For each player, the decisions of the other players constitute the relevant state of the world, which together with her decision, determines the consequences in accordance with the payoff matrix.

1. Start from the initial position, player i calculates expected utility and moves by her adaptive rule to a new state of indecision.

Games played by Bayesian deliberators

For each player, the decisions of the other players constitute the relevant state of the world, which together with her decision, determines the consequences in accordance with the payoff matrix.

1. Start from the initial position, player i calculates expected utility and moves by her adaptive rule to a new state of indecision.
2. She knows that the other players are Bayesian deliberators who have just carried out a similar process.

Games played by Bayesian deliberators

For each player, the decisions of the other players constitute the relevant state of the world, which together with her decision, determines the consequences in accordance with the payoff matrix.

1. Start from the initial position, player i calculates expected utility and moves by her adaptive rule to a new state of indecision.
2. She knows that the other players are Bayesian deliberators who have just carried out a similar process.
3. So, she can simply go through their calculations to see their new states of indecision and update her probabilities for their acts accordingly (*update by emulation*).

Games played by Bayesian deliberators

Under suitable conditions of common knowledge, a joint deliberational equilibrium on the part of all players corresponds to a Nash equilibrium point of the game.

Games played by Bayesian deliberators

Under suitable conditions of common knowledge, a joint deliberational equilibrium on the part of all players corresponds to a Nash equilibrium point of the game.

Strengthening the assumptions slightly leads in a natural way to refinements of the Nash equilibrium.

Games played by Bayesian deliberators

In a game played by Bayesian deliberators with a common prior, an adaptive rule that seeks the good, and a feedback process that updates by emulation, with common knowledge of all the foregoing, each players is at a deliberational equilibrium iff the corresponding mixed acts are a Nash equilibrium.

Games played by Bayesian deliberators

In a game played by Bayesian deliberators with a common prior, an adaptive rule that seeks the good, and a feedback process that updates by emulation, with common knowledge of all the foregoing, each players is at a deliberational equilibrium iff the corresponding mixed acts are a Nash equilibrium.

"mixed strategies as beliefs"

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

$\mathbf{P}_A = \langle 0.2, 0.8 \rangle$ and $\mathbf{P}_B = \langle 0.4, 0.6 \rangle$

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

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

$\mathbf{P}_A = \langle 0.2, 0.8 \rangle$ and $\mathbf{P}_B = \langle 0.4, 0.6 \rangle$

$$EU(U) = 0.8 \qquad COV(U) = \max(0.8 - 0.64, 0) = 0.16$$
$$EU(D) = 0.6 \qquad COV(D) = \max(0.6 - 0.64, 0) = 0$$
$$EU(L) = 0.2 \qquad COV(L) = \max(0.28 - 1.04, 0) = 0$$
$$EU(R) = 1.6 \qquad COV(R) = \max(1.6 - 1.04, 0) = 0.56$$
$$SQ_A = 0.64$$
$$SQ_B = 1.04$$

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

$\mathbf{P}_A = \langle 0.2, 0.8 \rangle$ and $\mathbf{P}_B = \langle 0.4, 0.6 \rangle$

$EU(U) = 0.8$ $\quad COV(U) = \max(0.8 - 0.64, 0) = 0.16$

$EU(D) = 0.6$ $\quad COV(D) = \max(0.6 - 0.64, 0) = 0$

$EU(L) = 0.2$ $\quad COV(L) = \max(0.28 - 1.04, 0) = 0$

$EU(R) = 1.6$ $\quad COV(R) = \max(1.6 - 1.04, 0) = 0.56$

$p_U = \frac{k \cdot 0.2 + 0.16}{k + 0.16}$

$p_L = \frac{k \cdot 0.4 + 0}{k + 0.56}$

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

$\mathbf{P}_A = \langle 0.2, 0.8 \rangle$ and $\mathbf{P}_B = \langle 0.4, 0.6 \rangle$

$$EU(U) = 0.8$$
$$EU(D) = 0.6$$
$$EU(L) = 0.2$$
$$EU(R) = 1.6$$
$$p_U = \frac{10 \cdot 0.2 + 0.16}{10 + 0.16} = 0.212598$$
$$p_L = \frac{k \cdot 0.4 + 0}{k + 0.56} = 0.378788$$

$$COV(U) = \max(0.8 - 0.64, 0) = 0.16$$
$$COV(D) = \max(0.6 - 0.64, 0) = 0$$
$$COV(L) = \max(0.28 - 1.04, 0) = 0$$
$$COV(R) = \max(1.6 - 1.04, 0) = 0.56$$

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

$\mathbf{P}_A = \langle 0.212598, 0.787402\rangle$ and $\mathbf{P}_B = \langle 0.378788, 0.621212\rangle$

$$EU(U) = 0.38 \cdot 2 + 0.62 \cdot 0 = 0.8$$
$$EU(D) = 0.38 \cdot 0 + 0.62 \cdot 1 = 0.6$$
$$EU(L) = 0.21 \cdot 1 + 0.78 \cdot 0 = 0.2$$
$$EU(R) = 0.21 \cdot 0 + 0.78 \cdot 2 = 1.6$$
$$SQ_A = 0.21 \cdot EU(U) + 0.78 \cdot EU(D)$$
$$SQ_B = 0.37 \cdot EU(L) + 0.62 \cdot EU(R)$$

Bayes Dynamics

If the new information that a player gets by emulating other players' calculations, updating his probabilities on their actions, and recalculating his expected utilities is e, then his new probabilities that he will do act A should be:

$$p_2(A) = p_1(A) \cdot \frac{p(e \mid A)}{\sum_i p(A_i) \cdot p(e \mid A_i)}$$

where $\{A_i\}$ is a partition on the alternative acts.

Bayes Dynamics

If the new information that a player gets by emulating other players' calculations, updating his probabilities on their actions, and recalculating his expected utilities is e, then his new probabilities that he will do act A should be:

$$p_2(A) = p_1(A) \cdot \frac{p(e \mid A)}{\sum_i p(A_i) \cdot p(e \mid A_i)}$$

where $\{A_i\}$ is a partition on the alternative acts.

But our deliberators do not have the appropriate proposition e in a large probability space that defines the likelihoods $p(e \mid A)$.

Is Nash a Bayes dynamics?

- If a deliberator starts with probability 1 that she will do some act that has utility less than the status quo, Nash will pull that probability down and raise the zero probabilities of competing acts.

Is Nash a Bayes dynamics?

- If a deliberator starts with probability 1 that she will do some act that has utility less than the status quo, Nash will pull that probability down and raise the zero probabilities of competing acts.

  "Indeed, one can argue that if a deliberator is absolutely sure which act he is going to do he needn't deliberate, and if he is absolutely sure he won't do an act, then his deliberation should ignore that act. "

Is Nash a Bayes dynamics?

- If a deliberator starts with probability 1 that she will do some act that has utility less than the status quo, Nash will pull that probability down and raise the zero probabilities of competing acts.

  "Indeed, one can argue that if a deliberator is absolutely sure which act he is going to do he needn't deliberate, and if he is absolutely sure he won't do an act, then his deliberation should ignore that act. "

- If two acts have expected utility less that the status quo, then they both get covetability 0, even if their expected utilities are quite different.

Tendency toward better response

The present expected utilities may not be the final ones, but they are the players' "best guess"

Assume that the decision makers likelihoods are an increasing function of the newly calculated expected utilities.

Tendency toward better response

The present expected utilities may not be the final ones, but they are the players' "best guess"

Assume that the decision makers likelihoods are an increasing function of the newly calculated expected utilities.

Darwin flow:

$$p_2(A) = k \cdot \frac{EU(A) - EU(SQ)}{EU(SQ)}$$

Refinements for Nash equilibrium

Bob

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

Refinements for Nash equilibrium

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

If Bayesian deliberation must start in the interior of the space of indecision, then *dynamic deliberation* cannot lead to U, R.

Refinements for Nash equilibrium

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

If Bayesian deliberation must start in the interior of the space of indecision, then *dynamic deliberation* cannot lead to U, R.

Call an equilibrium *accessible* provided one can converge to it starting at a completely mixed state of indecision.

Refinements for Nash equilibrium

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

If Bayesian deliberation must start in the interior of the space of indecision, then *dynamic deliberation* cannot lead to U, R.

Call an equilibrium *accessible* provided one can converge to it starting at a completely mixed state of indecision.

Does *accessibility* correspond to perfect/proper equilibria?

		Bob L	R
Ann	U	0.5,0.5	0.5,0.5
	D	1,1	0,0

		Bob	
		L	R
Ann	U	0.5,0.5	0.5,0.5
	D	1,1	0,0

Darwin can lead to an imperfect equilibrium. Nash can only lead to D, L.

		Bob	
		L	R
Ann	U	0.5,0.5	0.5,0.5
	D	1,1	0,0

Darwin can lead to an imperfect equilibrium. Nash can only lead to D, L.

Samuelson identified adaptive rules that correspond to proper/perfect equilibrium. A key feature is:

ordinality: the velocity of probability change of a strategy depends only on the ordinal ranking among strategies according to their expected utilities.

L. Samuelson. *Evolutionary foundations for solution concepts for finite, two-player, normal-form games*. Proceedings of TARK, 1988.

Coordination

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

Coordination

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

1. How can convention without communication be sustained? (Lewis)
2. How can convention without communication be generated?

Ann and Bob each have predeliberational probabilities. They can be anything at all. These probabilities are made common knowledge at the start of deliberation.

Ann and Bob each have predeliberational probabilities. They can be anything at all. These probabilities are made common knowledge at the start of deliberation.

You—the philosopher—have some probability distribution over the space of Ann and Bob's initial probabilities. Then you should believe with probability one that the deliberators will converge to one of the pure Nash equilibria.

Ann and Bob each have predeliberational probabilities. They can be anything at all. These probabilities are made common knowledge at the start of deliberation.

You—the philosopher—have some probability distribution over the space of Ann and Bob's initial probabilities. Then you should believe with probability one that the deliberators will converge to one of the pure Nash equilibria.

Precedent and other forms of initial salience may influence the deliberators' initial probabilities, and thus may play a role in determining *which* equilibrium is selected.

Ann and Bob each have predeliberational probabilities. They can be anything at all. These probabilities are made common knowledge at the start of deliberation.

You—the philosopher—have some probability distribution over the space of Ann and Bob's initial probabilities. Then you should believe with probability one that the deliberators will converge to one of the pure Nash equilibria.

Precedent and other forms of initial salience may influence the deliberators' initial probabilities, and thus may play a role in determining *which* equilibrium is selected.

Coordination is effected by rational deliberation.

Ann and Bob each have predeliberational probabilities. They can be anything at all. These probabilities are made common knowledge at the start of deliberation.

You—the philosopher—have some probability distribution over the space of Ann and Bob's initial probabilities. Then you should believe with probability one that the deliberators will converge to one of the pure Nash equilibria.

Precedent and other forms of initial salience may influence the deliberators' initial probabilities, and thus may play a role in determining *which* equilibrium is selected.

Coordination is effected by rational deliberation.

The answer to the question of how convention can be generated for Bayesian deliberators has both methodological and psychological aspects.

Stability

An equilibrium point e is **stable** under the dynamics if points nearby remain close for all time under the action of the dynamics. It is **strongly stable** if there is a neighborhood of e swuch that the trajectories of all points in that neighborhood converge to e.

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

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

- A dynamically unstable equilibrium is a natural focus of worry about trembling hands: confining the trembles to an arbitrary small neighborhood cannot guarantee that the trajectory stays "close by"

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

- A dynamically unstable equilibrium is a natural focus of worry about trembling hands: confining the trembles to an arbitrary small neighborhood cannot guarantee that the trajectory stays "close by"
- static vs. dynamic view of stability: in the static view, mixed strategies are not stable, but in the dynamic view strategies may or may not be stable.

General comments

- Extensive games, imprecise probabilities, other notions of stability, weaken common knowledge assumptions,...
- Generalizing the basic model
- Why assume deliberators are in a "information feedback situation"?
- Deliberation in decision theory.

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.

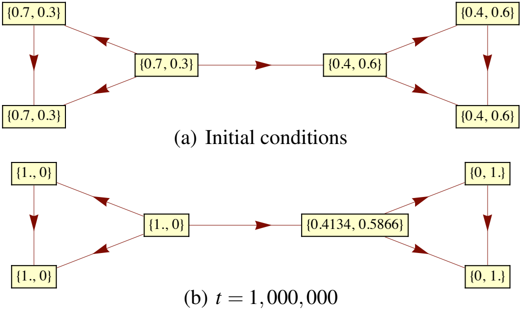

Fig. 8 Battle of the Sexes played by Nash deliberators ($k = 25$) on two cycles connected by a bridge edge (values rounded to the nearest 10^{-4}).

Tomorrow: Common modes of reasoning.