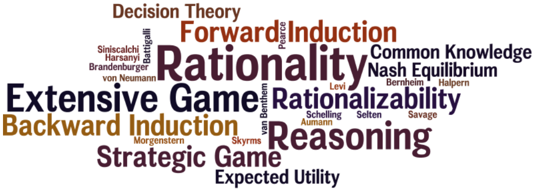
Reasoning in Games: Players as Programs

Lecture 5

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
- **Friday** *brief* introduction to webppl; Game-theoretic reasoning in webppl; Coordination games (comparing Skyrms's model of deliberation and the webppl approach); Models of game-theoretic reasoning

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

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

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: $SQ_i = \sum_{s_i \in S_i} I_i(s_i) EU_i(s_i)$.

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".

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 $s \in S_i$:

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

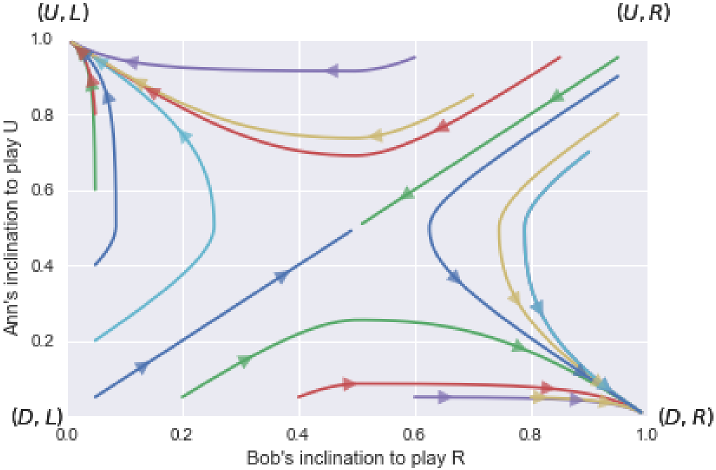
where $k > 0$ is the “index of caution".

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.

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

Coordination - Nash deliberators



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.

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

Correlated Strategies

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

- Three Nash equilibria:
  - (U, L): the payoff is $(2, 1)$
  - (D, R): the payoff is $(1, 2)$
  - $([\frac{2}{3}(U), \frac{1}{3}D], [\frac{1}{3}(L), \frac{2}{3}(R)])$: the payoff is $(\frac{2}{3}, \frac{2}{3})$

Correlated Strategies

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

- Three Nash equilibria:
  - (U, L): the payoff is $(2, 1)$
  - (D, R): the payoff is $(1, 2)$
  - $([\frac{2}{3}(U), \frac{1}{3}D], [\frac{1}{3}(L), \frac{2}{3}(R)])$: the payoff is $(\frac{2}{3}, \frac{2}{3})$
- Mixed Strategies: Each player conducts a private, independent lottery to choose their strategy.

Correlated Strategies

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

	L	R
U	0.5	0
D	0	0.5

- Three Nash equilibria:
  - (U, L): the payoff is $(2, 1)$
  - (D, R): the payoff is $(1, 2)$
  - $([\frac{2}{3}(U), \frac{1}{3}D], [\frac{1}{3}(L), \frac{2}{3}(R)])$: the payoff is $(\frac{2}{3}, \frac{2}{3})$
- Mixed Strategies: Each player conducts a private, independent lottery to choose their strategy.
- Conduct a *public* lottery: flip a fair coin and follow the strategy ($H \Rightarrow (U, L)$, $T \Rightarrow (D, R)$). The payoff is $(1.5, 1.5)$.

Two extremes:

1. Completely private, independent lotteries
2. A single, completely public lottery

Two extremes:

1. Completely private, independent lotteries
2. A single, completely public lottery

What about: a public lottery, but reveal only partial information about the outcome to each of the players?

Correlation

Correlation: Players can improve their expected value by correlating their choices on an “outside signal"

Correlation

Correlation: Players can improve their expected value by correlating their choices on an “outside signal"

With more than 2 players...

- A player may believe that (some of) the other players strategy choices are **independent** or **correlated**.
- Two players can **agree** or **disagree** on the probabilities that the assign to a third player’s choice of strategy.

Characterizing Correlated Equilibrium

Theorem (Aumann). σ is a correlated equilibrium of G iff there exists a model $\mathcal{M}^G = \langle W, (P_i)_{i\in N}, \mathbf{s}\rangle$ such that:

- for all $i \in N$, $\mathsf{Rat}_i = W$; and
- for all $i \in N$, $P_i^S = \sigma$.

Imagine an outside observer, who does not know what the players' initial probabilities for the possible actions will be, but rather has his own probability measure over the possible initial states of indecision of the system.

With respect to this probability, the players are at a correlated equilibrium.

Imagine an outside observer, who does not know what the players' initial probabilities for the possible actions will be, but rather has his own probability measure over the possible initial states of indecision of the system.

With respect to this probability, the players are at a correlated equilibrium.

"This correlated equilibrium is a general result of the players' common knowledge and Bayesian dynamic deliberations." (Skyrms, pg. 60)

The same result may be obtained without the outside observer if prior to delibreration the players themselves share the role of the outside observer.

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

The WebPPL language is a subset of JavaScript with extra syntax to describe *probabilistic computation*.

E.g., there is are primitive operations that describe deterministic functions (such as *and*) and stochastic operations, such as the *flip function*.

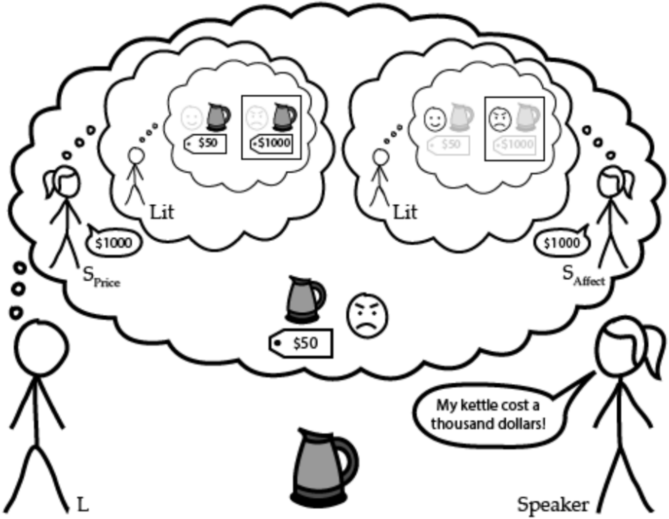
The WebPPL language is a subset of JavaScript with extra syntax to describe *probabilistic computation*.

E.g., there is are primitive operations that describe deterministic functions (such as *and*) and stochastic operations, such as the *flip function*.

Probabilistic computation:

$f_T : \{0,1\}^* \to \{0,1\}^*$
(T is a Turing machine)

$f_T : \{0,1\}^* \to \Delta(\{0,1\}^*)$
(T a probabilistic Turing machine)



N. D. Goodman and M. C. Frank. *Pragmatic language interpretation as probabilistic inference*. Trends in Cognitive Sciences, 2016.

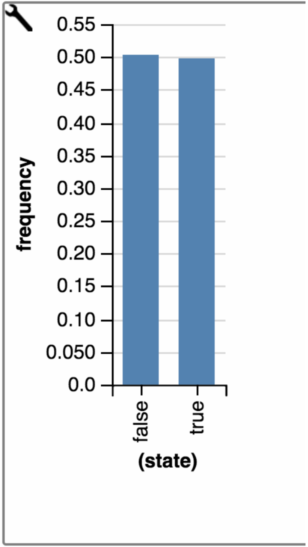
probmods.org

Listing 1 Introduction to webppl

```
//flip a fair coin
flip()

// visualize coin flips
viz(repeat(1000,flip))
```

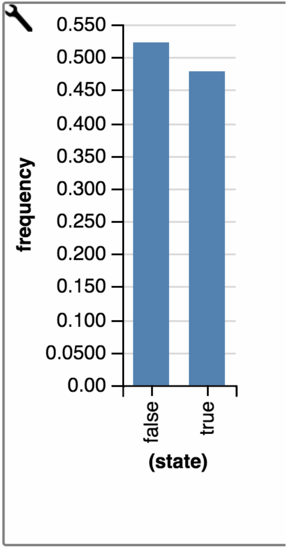
Listing 2 Introduction to webppl

```
//flip a fair coin
flip()

// visualize coin flips
viz(repeat(1000,flip))
```



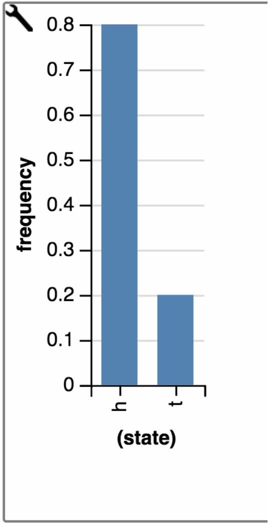
Listing 3 Introduction to webppl

```
//flip a fair coin
flip()

// visualize coin flips
viz(repeat(1000,flip))
```



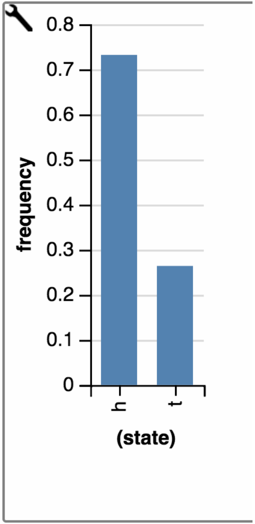
Listing 4 Introduction to webppl

```
var trickCoin = function() {
    flip(0.75) ? 'h' : 't' };

viz(repeat(100, trickCoin))
```



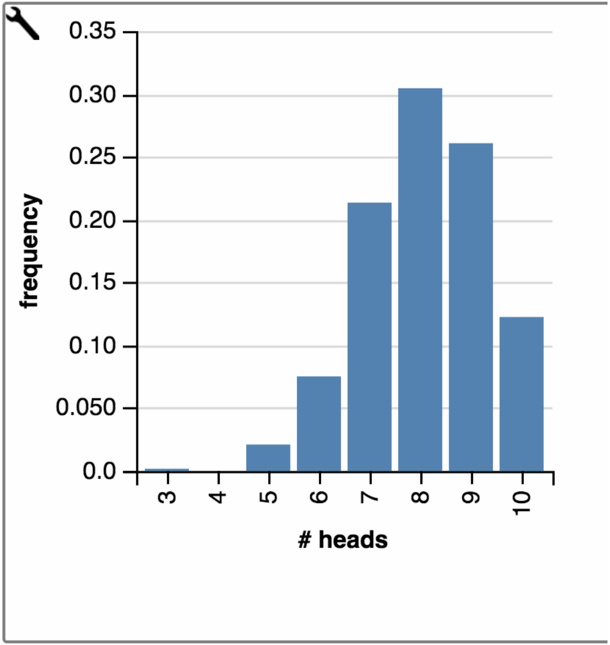
Listing 5 Introduction to webppl

```
var trickCoin = function() {
    flip(0.75) ? 'h' : 't' };

viz(repeat(1000, trickCoin))
```



Listing 6 Introduction to webppl

```
var makeCoin = function(weight) {
    return function() { return flip(weight) }
}

var coin = makeCoin(0.8)

var data = repeat(1000, function() { sum(repeat(10, coin)) })

viz(data, {xLabel: '# heads'})
```



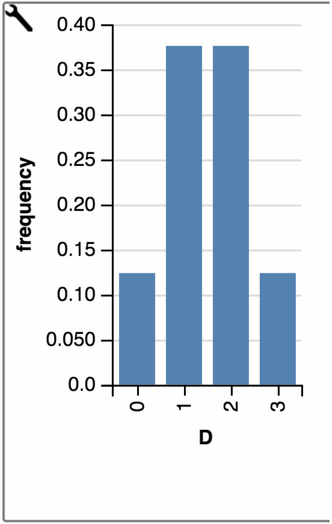
0.35
0.30
0.25
0.20
0.15
0.10
0.050
0.0
frequency
3
4
5
6
7
8
9
10
heads

Listing 7 Introduction to webppl

```
var model = function() {
    var A = flip()
    var B = flip()
    var C = flip()
    var D = A + B + C
    return {"D": D}
  }
  var dist = Infer({}, model)
  viz(dist)
```

Listing 8 Introduction to webppl

```
var model = function() {
    var A = flip()
    var B = flip()
    var C = flip()
    var D = A + B + C
    return {"D": D}
  }
  var dist = Infer({}, model)
  viz(dist)
```

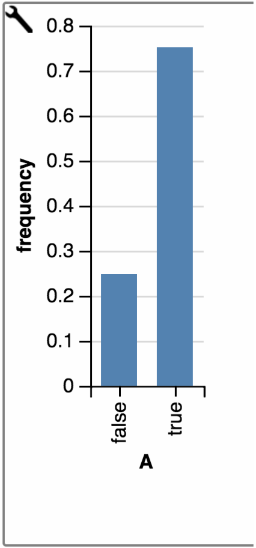


Listing 9 Introduction to webppl

```
var model = function () {
    var A = flip()
    var B = flip()
    var C = flip()
    var D = A + B + C

    condition(D >= 2)
    return {'A': A}
  };
  var dist = Infer({}, model)
  viz(dist)
```

Listing 10 Introduction to webppl

```
var model = function () {
    var A = flip()
    var B = flip()
    var C = flip()
    var D = A + B + C

    condition(D >= 2)
    return {'A': A}
  };
  var dist = Infer({}, model)
  viz(dist)
```

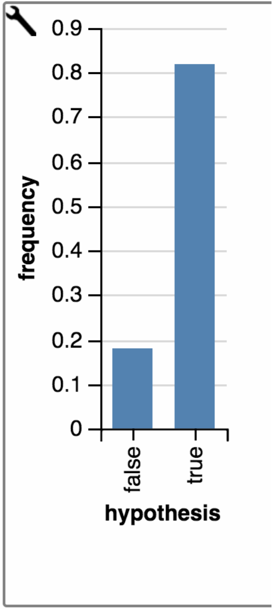


Bayes Theorem

$$Pr(H \mid D) = \frac{Pr(D \mid H)Pr(H)}{Pr(D)}$$

Listing 11 Bayes Theorem

```
var observedData = true;
var prior = function () { flip(.6) }
var likelihood = function (h) { h ? flip(0.75) : flip(0.25) }

var posterior = Infer({method: "enumerate"},
  function () {
    var hypothesis = prior()
    var data = likelihood(hypothesis)
    condition(data == observedData)
    return {hypothesis: hypothesis}
})

viz(posterior)
```


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
frequency
false
true
hypothesis

Coordination in WebPPL

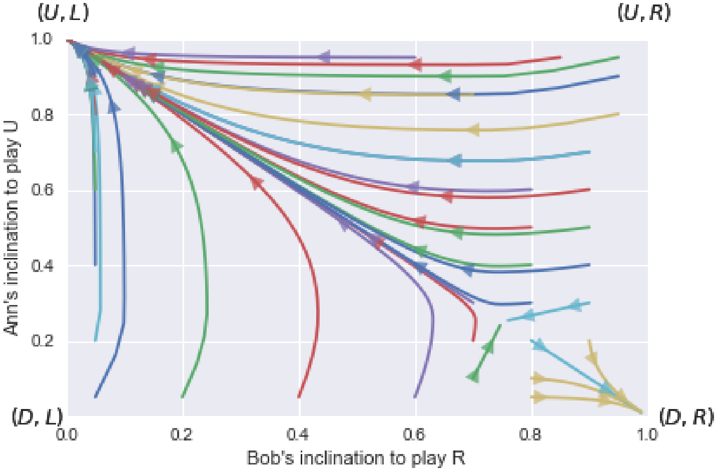
agentmodels.org/chapters/7-multi-agent.html

A. StuhlmÏler and N. D. Goodman. *Reasoning about Reasoning by Nested Conditioning: Modeling Theory of Mind with Probabilistic Programs*. Journal Cognitive Systems Research, 2014.

Hi-Low

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

Hi-lo game - Nash deliberators



Focal Points

> *"There are these two broad empirical facts about Hi-Lo games, people almost always choose A [Hi] and people with common knowledge of each other's rationality think it is obviously rational to choose A [Hi]."*
>
> [Bacharach, *Beyond Individual Choice*, 2006, pg. 42]

See also chapter 2 of:

C.F. Camerer. *Behavioral Game Theory*. Princeton UP, 2003.

N. Bardsley, J. Mehta, C. Starmer and R. Sugden. *The Nature of Salience Revisited: Cognitive Hierarchy Theory* versus *Team Reasoning*. Economic Journal.

‘primary salience’: players’ psychological propensities to play particular strategies by default, when there are no other reasons for choice.

‘primary salience’: players’ psychological propensities to play particular strategies by default, when there are no other reasons for choice.

(*level-*n theory/ *cognitive hierarchy* theory)

‘primary salience’: players’ psychological propensities to play particular strategies by default, when there are no other reasons for choice.

(*level-n* theory/ *cognitive hierarchy* theory)

‘team reasoning’: assumes that each player chooses the decision rule which, if used by all players, would be optimal for each of them.

Do the two approaches make different predictions?

What do the experiments support?

pickers: choose between labels without any incentive to choose one rather than the other

pickers: choose between labels without any incentive to choose one rather than the other

guessers: guess how pickers have behaved

pickers: choose between labels without any incentive to choose one rather than the other

guessers: guess how pickers have behaved

coordinators: try to coordinate their choices

pickers: choose between labels without any incentive to choose one rather than the other

guessers: guess how pickers have behaved

coordinators: try to coordinate their choices

labels vs. *options*

$$\{water, beer, sherry, whisky, wine\}$$

$$\{water, beer, sherry, whisky, wine\}$$

Task 1: pick an option

$$\{\mathbf{water}, \mathit{beer}, \mathit{sherry}, \mathit{whisky}, \mathit{wine}\}$$

Task 1: pick an option

$$\{\mathbf{water}, \mathit{beer}, \mathit{sherry}, \mathit{whisky}, \mathit{wine}\}$$

Task 1: pick an option
Task 2: guess what your opponent picked

$$\{\mathbf{water}, \mathit{beer}, \mathit{sherry}, \mathit{whisky}, \mathit{wine}\}$$

Task 1: pick an option
Task 2: guess what your opponent picked
Task 3: try to coordinate with your (unknown) partner

$\{\mathbf{water}, \textit{beer}, \textit{sherry}, \textit{whisky}, \textit{wine}\}$

Task 1: pick an option
Task 2: guess what your opponent picked
Task 3: try to coordinate with your (unknown) partner

	pick	guess	coordinate
water	20	15	38
beer	13	26	11
sherry	4	1	0
whisky	6	6	5
wine	10	4	2

“The main aim of the two experiments was to test cognitive hierarchy theory and the theory of team reasoning as rival explanations of behaviour in pure coordination and Hi-Lo games.

“The main aim of the two experiments was to test cognitive hierarchy theory and the theory of team reasoning as rival explanations of behaviour in pure coordination and Hi-Lo games. Formally, our conclusion must be that each theory failed at least one test."

“The main aim of the two experiments was to test cognitive hierarchy theory and the theory of team reasoning as rival explanations of behaviour in pure coordination and Hi-Lo games. Formally, our conclusion must be that each theory failed at least one test."

“ The implication is that our subjects were able to use subtle features of the experimental environment to solve the problem of coordinating on a common mode of reasoning. This behaviour reveals an ability to solve coordination problems at a conceptual level above that of the theories of cognitive hierarchy and team reasoning that we have been examining. Each of those theories captures certain aspects of focal-point reasoning, but some essential feature of the human ability to solve coordination problems seems to have escaped formalisation."

“The basic intellectual premise, or working hypothesis, for rational players in this game seems to be the premise that some rule must be used if success is to exceed coincidence, and that the best rule to be found, whatever its rationalization, is consequently a rational rule." (Thomas Schelling)

Concluding remarks

Have we captured *strategic reasoning*?

Strategic reasoning vs. Bayesian rationality

- Normal form vs. Extensive Form: Should the analysis take place on the tree or the matrix? (plans vs. strategies)
- There is an important different between what I would believe given E is true and what I believe after *learning* E
- What should I assume about my opponents?
- What is the role of *higher-order beliefs*? (Common knowledge, common belief)
- Framing issues/language in game theory
- $\cdots$

“...[W]e cannot expect game and economic theory to be descriptive in the same sense that physics or astronomy are. Rationality is only one of several factors affecting human behavior; no theory based on this one factor alone can be expected to yield reliable predictions.

In fact, I find it somewhat surprising that our disciplines have any relation at all to real behavior. (I hope that most readers will agree that there is indeed such a relation, that we do gain some insight into the behavior of *Homo sapiens* by studying *Homo rationalis*.)"

R. Aumann. *What is game theory trying to accomplish?*. Frontiers of Economics, 1985.

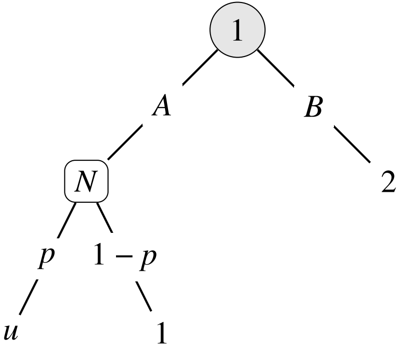
Can a player assign subjective probabilities to strategies under the control of other players who have their own objectives?

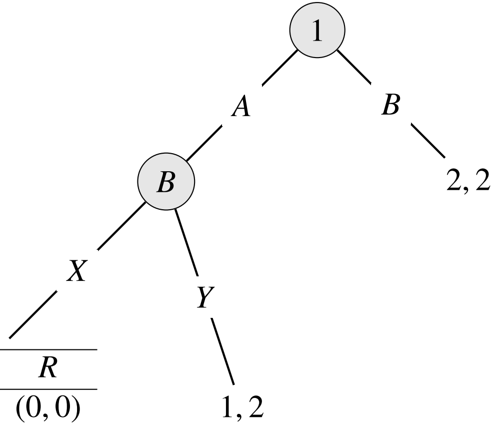
M. Mariotti. *Is Bayesian Rationality Compatible with Strategic Rationality?*. The Economic Journal, 105: 432, pgs. 1099 - 1109, 1995.

M. Mariotti. *Decisions in games: why there should be a special exemption from Bayesian rationality*. Journal of Economic Methodology, 4: 1, pgs. 43 - 60, 1997.

P. Hammond. *Expected Utility in Non-Cooperative Game Theory*. in *Handbook of Utility Theory*, 2004.

Games as consequences: “A decision maker prefers to be player i in game G_1 to being player j in game G_2"


1
A
B
N
2
p
1 − p
u
1



	L	R
U	(1, 7)	(0, 0)
D	(0, 0)	(3, 3)

Can the decision problem be *separated* from the game situation?

Can the decision problem be *separated* from the game situation?

Are strategies merely neutral access routes to consequences?

E. McClennen. *Rational choice in the context of ideal games.* in *Knowledge, Belief and Strategic Interaction*, pgs. 47-60, 1992.

utility must be measured *in the context of the game itself*.

I. Gilboa and D. Schmeidler. *A Derivation of Expected Utility Maximization in the Context of a Game*. Games and Economic Behavior, 44, pgs. 184 - 194, 2003.

The following two outcomes are not equivalent:

- “I get $90"
- “I get $90 and choose to leave $10 to my opponent"

The following two outcomes are not equivalent:

- “I get $10 and player one gets $90, and this was decided by Nature"
- “I get $10, player one gets $90 and this was decided by Player one".

Players need two theories:

1. A theory to guide their decisions.
2. A theory to predict the behavior of their opponents.

“Game theory is decision theory about special decision makers, namely about decision makers who theorize decision-theoretically about the other persons figuring in their decision situations.” (Spohn, “How to make sense of Game Theory")

“Rationality has a clear interpretation in individual decision making, but it does not transfer comfortably to interactive decisions, because interactive decision makers cannot maximize expected utility without strong assumptions about how the other participant(s) will behave. In game theory, common knowledge and rationality assumptions have therefore been introduced, but under these assumptions, rationality does not appear to be characteristic of social interaction in general." (pg. 152)

A. Colman. *Cooperation, psychological game theory, and limitations of rationality in social interaction*. Behavioral and Brain Sciences, 26, pgs. 139 - 198, 2003.

		Bob	
		L	R
Ann	U	3,-3	-1,1
	D	-9,9	3,-3

What is your advice to Ann for optimal play? Should she play U or D on the next move and how should he decide?

What is your advice to Ann for optimal play? Should she play U or D on the next move and how should he decide? Do you have a general strategy to recommend to Ann for the next 100 plays? If so, is your recommended strategy independent of or conditional on Bob's behavior during the 100-play sequence?

What is your advice to Ann for optimal play? Should she play U or D on the next move and how should he decide? Do you have a general strategy to recommend to Ann for the next 100 plays? If so, is your recommended strategy independent of or conditional on Bob's behavior during the 100-play sequence? What literatures would you draw upon for your advice to Ann in this situation?

What is your advice to Ann for optimal play? Should she play U or D on the next move and how should he decide? Do you have a general strategy to recommend to Ann for the next 100 plays? If so, is your recommended strategy independent of or conditional on Bob's behavior during the 100-play sequence? What literatures would you draw upon for your advice to Ann in this situation? Game theory? Classical statistical decision theory? Bayesian decision theory? Psychology?

J. Kadane and P. Larkey. *The Confusion of Is and Ought in Game Theoretic Contexts*. Management Sciences, 29:12, pgs. 1365 - 1379.

		Bob	
		L	R
Ann	U	1,0	3,2
	D	2,1	4,0

Ann \ Bob	L	R
U	1,0	3,2
D	2,1	4,0

		Bob	
		L	R
Ann	U	1,0	3,2
	D	2,1	4,0

Can Ann “teach" Bob to play R?

Weaken the common knowledge assumptions (payoffs, beliefs, dynamical rule, updating by emulation)

Chapter 6 of the Skyrms book is on *good habits*:

“A good habit is one which can be expected to save more in costs of reasoning than it is expected to lose by foregoing an extensive analysis of the decision involved....I explore the possibility that decisionmakers may be even more bounded in that they may find it economical to substitute special-purpose habits of direct strategy selections for Bayesian deliberation. "

(Skyrms, pg. 126)

“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)

Thank You!

pacuit.org

epacuit@umd.edu