Reasoning, Games, Action and Rationality

Lecture 1

ESSLLI'08, Hamburg

Eric Pacuit — Olivier Roy

Stanford University — University of Groningen

ai.stanford.edu/~epacuit — philos.rug.nl/~epacuit

February 20, 2009

Plan for Today

- General Motivations : Bayesian Rationality
- Games
- Models of Information in Games. Type Structures and Epistemic Plausibility models
- Rationality

Bayesian Rationality

- Instrumental Rationality
- Decision Theory
  - Endogenous and Exogenous Uncertainty
  - Maximization of Expected Utility

...to understand the fundamental ideas of game theory, one should begin by studying decision theory. -R. Myerson (*Game Theory*)

Just Enough Game Theory

> *"Game theory is a bag of analytical tools designed to help us understand the phenomena that we observe when decision-makers interact."*

Osborne and Rubinstein. *Introduction to Game Theory*. MIT Press .

Just Enough Game Theory

> *"Game theory is a bag of analytical tools designed to help us understand the phenomena that we observe when decision-makers interact."*

Osborne and Rubinstein. *Introduction to Game Theory*. MIT Press .

A **game** is a description of strategic interaction that includes

- actions the players *can* take
- description of the players' interests (i.e., preferences),

Just Enough Game Theory

> *"Game theory is a bag of analytical tools designed to help us understand the phenomena that we observe when decision-makers interact."*

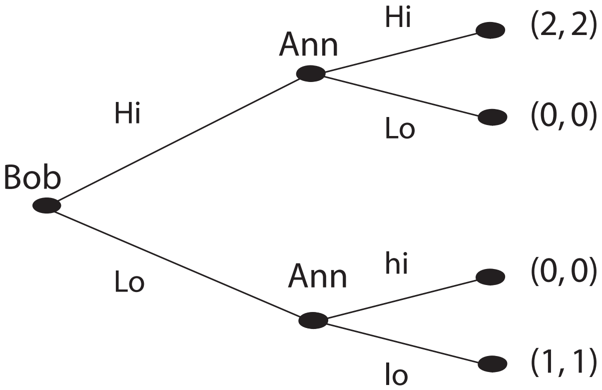
Osborne and Rubinstein. *Introduction to Game Theory*. MIT Press .

A **game** is a description of strategic interaction that includes

- actions the players *can* take
- description of the players' interests (i.e., preferences),

It does not specify the actions that the players do take.

Situations of Interaction - Games in Extensive Forms



Definition

Strategic games A *strategic game* $\mathbb{G}$ is a tuple $\langle I, S_i, v_i \rangle$ such that :

- I is a finite set of agents.
- S_i is a finite set of *actions* or *strategies* for i. A *strategy profile* $\sigma \in \Pi_{i \in I} S_i$ is a vector of strategies, one for each agent in I. The strategy s_i which i plays in the profile σ is noted σ_i.
- $v_i : \Pi_{i \in I} S_i \longrightarrow \mathbb{R}$ is an *utility function* that assigns to every strategy profile $\sigma \in \Pi_{i \in I} S_i$ the utility valuation of that profile for agent i.

Situations of Interaction

	A	B
a	1, 1	0, 0
b	0, 0	1, 1

Situations of Interaction

	A	B
a	1, 1	0, 0
b	0, 0	1, 1

What is the rational thing to do for Ann?

- It depends on what she *expects* Bob to do.
- But this depends on what she thinks Bob expects her to do.
- And so on...

Information in games

Formally, a game is described by its strategy sets and payoff functions.

Information in games

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game.

Information in games

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12 % respectively.

Information in games

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12 % respectively. But the political situations are quite different.

Information in games

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties hold a third of the seats, or, say, 49%, 39%, and 12 % respectively. But the political situations are quite different. The difference lies in the attitudes of the players, in their expectations about each other, in custom, and in history, though the rules of the game do not distinguish between the two situations.

R. Aumann and J. H. Dreze. *When all is said and done, how should you play and what should you expect?*. Center for the Study of Rationality, 2005.

Information in games

- Various states of information disclosure.

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*
- Various models of information:

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*
- Various models of information:
  - Type structures and Epistemic plausibility models.

Harsanyi Type Space

A (game-theoretic) **type** of a player summarizes everything the player knows privately at the beginning of the game which could affect his beliefs about payoffs in the game and about all other players' types.

(Harsanyi argued that all uncertainty in a game can be equivalently modeled as uncertainty about payoff functions.)

Harsanyi Type Space:

$$\mathcal{T} = \langle \mathcal{A}, S, \{T_i\}_{\in \mathcal{A}}, \{\lambda_i\}_{i \in \mathcal{A}} \rangle$$

- $\mathcal{A}$ is a finite set of n agents
- S is the uncertainty domain
- T_i is a set of types
- $\lambda_i : T_i \rightarrow \Delta(S \times T_{-i})$

A state of the world is a tuple

$$(s, t_1, \ldots, t_n) \in S \times T_1 \times \cdots \times T_n$$

Epistemic Plausibility Model

Definition

An *epistemic plausibility model* $\mathbb{M}$ of the game $\mathbb{G}$ is a tuple $\langle W, f, \{\leq_i, \sim_i\}_{i \in I}\rangle$ such that:

- W is a set of *states*.
- Then $f : W \longrightarrow \Pi_{i \in I} S_i$ is a *strategy function* that assigns to each $w \in W$ a strategy profile. From convenience we write $\sigma(w)$ for the $\sigma = f(w)$ and $\sigma_i(w)$ for the i^{th} component of this profile.
- $\sim_i$ is an *epistemic accessibility equivalence relation* such that if $w \sim_i w'$ then $\sigma_i(w) = \sigma_i(w')$. We write $[w]_i$ for $\{w' : w \sim_i w'\}$.
- $\leq_i$ is a reflexive and transitive *plausibility ordering* on W such that if $w \leq_i w'$ then $w \sim_i w'$. This relation is say to be *locally connected* when, for all w and $w', w'' \in [w]_i$, either $w' \leq_i w''$ or $w'' \leq_i w'$

An example

	A	B
a	1, 1	0, 0
b	0, 0	1, 1

An example

$\lambda_{Ann}(t_{Ann})$	**A**	**B**
u_{Bob}	1/2	0
T_{Bob}	1/2	0

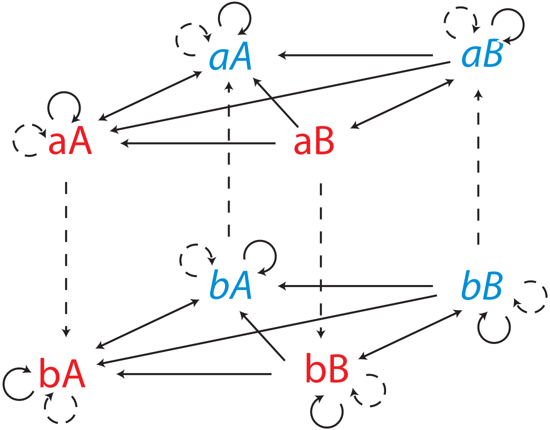
Table: Ann's beliefs about Bob

$\lambda_{Bob}(t_{Bob})$	**a**	**b**
u_{Ann}	1	0

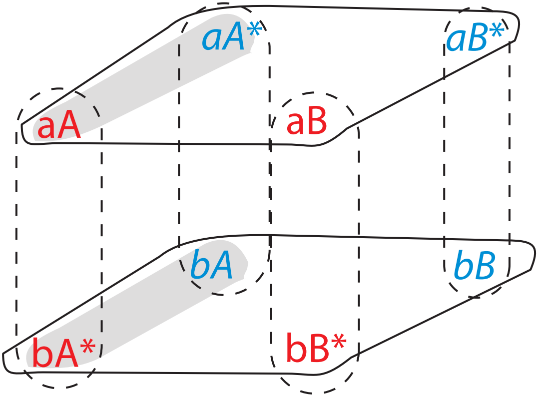
$\lambda_{Bob}(u_{Bob})$	**a**	**b**
u_{Ann}	0	1

Table: Bob's beliefs about Ann

An example



An example



Definition

Expected Value in type structure The *expected value* for player i of playing strategy s_i given that he is of type t_i is defined as follows.

$$EV_{t_i}(s_i) = \sum_{t'_{-i}} \sum_{\sigma'_{-i}} \lambda_i(t_i)(\sigma'_{-i}, t'_{-i}) v_i(s_i, \sigma'_{-i})$$

Definition

Expected Value in type structure The *expected value* for player i of playing strategy s_i given that he is of type t_i is defined as follows.

$$EV_{t_i}(s_i) = \sum_{t'_{-i}} \sum_{\sigma'_{-i}} \lambda_i(t_i)(\sigma'_{-i}, t'_{-i}) v_i(s_i, \sigma'_{-i})$$

Definition

Rationality in Epistemic plausibility models Given a state w, we write $w[s_i/w_i]$ for the profile σ that is just like $\sigma(w)$ except that $\sigma_i = s_i$. Player i is *irrational* at w when there is a $s'_i \neq \sigma_i(w)$ such that, $v_i(f(w')) \leq v_i(f(w'[s'_i/w'_i]))$ for all $w' \in max_{\leq_i}[w]_i$. Player i is *rational* at a state w when he is not irrational at that state.

Back to our example

$\lambda_{Ann}(t_{Ann})$	**A**	**B**
u_{Bob}	$1/2$	0
T_{Bob}	$1/2$	0

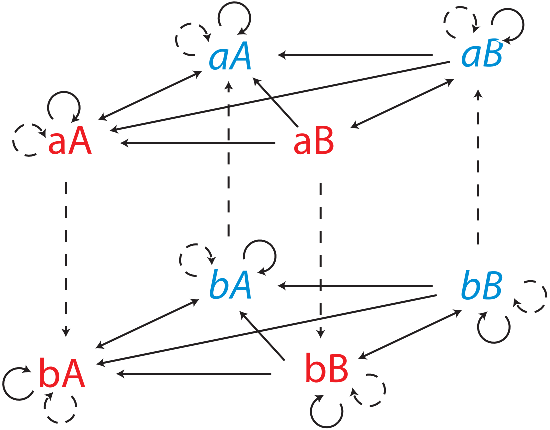
Table: Ann's beliefs about Bob

$\lambda_{Bob}(t_{Bob})$	**a**	**b**
u_{Ann}	1	0

$\lambda_{Bob}(u_{Bob})$	**a**	**b**
u_{Ann}	0	1

Table: Bob's beliefs about Ann

Back to our example



Tomorrow

- Logics to talk about these structures.
- Common knowledge of rationality and elimination of strictly dominati