Reasoning with Probabilities

Eric Pacuit Joshua Sack

September 11, 2009

Plan for the Course

✓ Introduction and Background

✓ Probabilistic Epistemic Logics

✓: Dynamic Probabilistic Epistemic Logics

Day 4: Reasoning with Probabilities

Day 5: Conclusions and General Issues

Plan for Today

- Reasoning with probabilities in games:
  1. Harsanyi Type Spaces
  2. Rationality
- Logics for Type Spaces

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are rational,

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are rational, each thinks each other is rational,

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are rational, each thinks each other is rational, each thinks each other thinks the others are rational,

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are rational, each thinks each other is rational, each thinks each other thinks the others are rational, and so on?

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are rational, each thinks each other is rational, each thinks each other thinks the others are rational, and so on?

Fundamental Question: What does it mean to say that the players in a strategic interactive situation are rational, each thinks each other is rational, each thinks each other thinks the others are rational, and so on?

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.

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.

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.

A **solution concept** is a systematic description of the outcomes that may emerge in a family of games.

This is the starting point for most of game theory and includes many variants: Nash equilibrium, backwards inductions, or iterated dominance of various kinds.

These are usually thought of as the embodiment of "rational behavior" in some way and used to analyze game situations.

Here, **solution concepts** are more of an *endpoint*.

A **solution concept** is a systematic description of the outcomes that may emerge in a family of games.

This is the starting point for most of game theory and includes many variants: Nash equilibrium, backwards inductions, or iterated dominance of various kinds.

These are usually thought of as the embodiment of "rational behavior" in some way and used to analyze game situations.

Here, **solution concepts** are more of an *endpoint*.

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

What does it mean for Ann to **be rational**? What is the rational thing for Ann to do?

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

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

What does it mean for Ann to **be rational**? What is the rational thing for Ann to do?

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

To answer these questions, we need a (mathematical) framework to study each of the following issues:

- Rationality: "Ann is rational"
- Knowledge/Beliefs: "Bob believes (knows) Ann is rational"
- Higher-order Knowledge/Beliefs: "Ann knows that Bob knows that Ann is rational", "it is common knowledge that all agents are rational".

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*
- Various "types" of information:
  - (hard information) own preferences, own beliefs, structure of the game, (soft information) what the other agent will do, etc.

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*
- Various "types" of information:
  - (hard information) own preferences, own beliefs, structure of the game, (soft information) what the other agent will do, etc.

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*
- Various "types" of information:
  - (hard information) own preferences, own beliefs, structure of the game, (soft information) what the other agent will do, etc.

Information in games

- Various states of information disclosure.
  - *Ex ante*, *ex interim*, *ex post*
- Various "types" of information:
  - (hard information) own preferences, own beliefs, structure of the game, (soft information) what the other agent will do, etc.

Describing the Players Knowledge and Beliefs

Fix a set of possible states (**complete descriptions of a situation**). Two main approaches to describe beliefs (knowledge):

- Set-theortical (Kripke Structures, Aumann Structures): For each state and each agent i, specify a set of states that i considers possible.
- Probabilistic (Bayesian Models, Harsanyi Type Spaces): For each state, define a (subjective) probability function over the set of states for each agent.

Harsanyi Type Space

John C. Harsanyi, nobel prize winner in economics, developed a theory of games with **incomplete information**.

J. Harsanyi. *Games with incomplete information played by "bayesian" players I-III. Management Science Theory* **14***: 159-182, 1967-68.*

Harsanyi Type Space

John C. Harsanyi, nobel prize winner in economics, developed a theory of games with **incomplete information**.

1. incomplete information: uncertainty about the *structure* of the game (outcomes, payoffs, strategy space)
2. imperfect information: uncertainty *within the game* about the previous moves of the players

J. Harsanyi. *Games with incomplete information played by "bayesian" players I-III. Management Science Theory* **14**: *159-182, 1967-68.*

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi Type Space

The problem: A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

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: The Basic Model

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

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

A state of the world is a tuple

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

Example

$$T_1 = \{t_1, t_1'\},\ T_2 = \{t_2, t_2'\},\ S = \{a, b\}$$

Player 1: $\lambda_1(t_1)$

	a	b
t_2	1	0
t_2'	0	0

$\lambda_1(t_1')$

	a	b
t_2	0	0
t_2'	0.3	0.7

Player 2: $\lambda_2(t_2)$

	a	b
t_1	0	0.5
t_1'	0.5	0

$\lambda_2(t_2')$

	a	b
t_1	0	0
t_1'	0	1

Example

$$T_1 = \{t_1, t_1'\},\ T_2 = \{t_2, t_2'\},\ S = \{a, b\}$$

Player 1: $\lambda_1(t_1)$

	a	b
t_2	1	0
t_2'	0	0

$\lambda_1(t_1')$

	a	b
t_2	0	0
t_2'	0.3	0.7

Player 2: $\lambda_2(t_2)$

	a	b
t_1	0	0.5
t_1'	0.5	0

$\lambda_2(t_2')$

	a	b
t_1	0	0
t_1'	0	1

t_1 is **certain** the outcome is a $(o = a)$.

Example

$$T_1 = \{t_1, t_1'\},\ T_2 = \{t_2, t_2'\},\ S = \{a, b\}$$

Player 1: $\lambda_1(t_1)$

	a	b
t_2	1	0
t_2'	0	0

$\lambda_1(t_1')$

	a	b
t_2	0	0
t_2'	0.3	0.7

Player 2: $\lambda_2(t_2)$

	a	b
t_1	0	0.5
t_1'	0.5	0

$\lambda_2(t_2')$

	a	b
t_1	0	0
t_1'	0	1

t_2 assigns probability 0.5 to player 1 being **certain** $o = a$.

Example

$$T_1 = \{t_1, t_1'\},\ T_2 = \{t_2, t_2'\},\ S = \{a, b\}$$

Player 1: $\lambda_1(t_1)$

	a	b
t_2	1	0
t_2'	0	0

$\lambda_1(t_1')$

	a	b
t_2	0	0
t_2'	0.3	0.7

Player 2: $\lambda_2(t_2)$

	a	b
t_1	0	0.5
t_1'	0.5	0

$\lambda_2(t_2')$

	a	b
t_1	0	0
t_1'	0	1

t_2' is **certain** player 1 is certain that he is certain the $o = b$.

Example

$$T_1 = \{t_1, t_1'\},\ T_2 = \{t_2, t_2'\},\ S = \{a, b\}$$

Player 1: $\lambda_1(t_1)$

	a	b
t_2	1	0
t_2'	0	0

$\lambda_1(t_1')$

	a	b
t_2	0	0
t_2'	0.3	0.7

Player 2: $\lambda_2(t_2)$

	a	b
t_1	0	0.5
t_1'	0.5	0

$\lambda_2(t_2')$

	a	b
t_1	0	0
t_1'	0	1

t_2' is **certain** player 1 is certain that he is certain that $o = b$.

More on Types

For simplicity, we assume $S = \times_{i \in \mathcal{A}} S_i$, where each S_i is a strategy space for agent i in some fixed game G. In this case, $\lambda_i : T_i \rightarrow \Delta(S_{-i} \times T_{-i})$.

A fixed state $(s_1, t_1, s_2, t_2, \ldots, s_n, t_n)$ specifies the strategies and each player's *entire hierarchy of beliefs*:

1. i's first-order beliefs: $T_i \mapsto \Delta(S_{-i} \times T_{-i}) \mapsto \Delta(S_{-i})$ (marginalizing)
2. i's second-order beliefs: $T_i \mapsto \Delta(S_{-i} \times T_{-i}) \mapsto \Delta(S^{-i} \times \times_{i \neq j} \Delta(S_{-j} \times T_{-j})) \mapsto \Delta(S_{-i} \times \times_{j \neq i} \Delta(S_{-j}))$ (marginalizing)

More on Types

For simplicity, we assume $S = \times_{i \in \mathcal{A}} S_i$, where each S_i is a strategy space for agent i in some fixed game G. In this case, $\lambda_i : T_i \to \Delta(S_{-i} \times T_{-i})$.

A fixed state $(s_1, t_1, s_2, t_2, \ldots, s_n, t_n)$ specifies the strategies and each player's *entire hierarchy of beliefs*:

1. i's first-order beliefs: $T_i \mapsto \Delta(S_{-i} \times T_{-i}) \mapsto \Delta(S_{-i})$ (marginalizing)
2. i's second-order beliefs: $T_i \mapsto \Delta(S_{-i} \times T_{-i}) \mapsto \Delta(S^{-i} \times \times_{i \neq j} \Delta(S_{-j} \times T_{-j})) \mapsto \Delta(S_{-i} \times \times_{j \neq i} \Delta(S_{-j}))$ (marginalizing)

More on Types

For simplicity, we assume $S = \times_{i \in \mathcal{A}} S_i$, where each S_i is a strategy space for agent i in some fixed game G. In this case, $\lambda_i : T_i \to \Delta(S_{-i} \times T_{-i})$.

A fixed state $(s_1, t_1, s_2, t_2, \ldots, s_n, t_n)$ specifies the strategies and each player's *entire hierarchy of beliefs*:

1. i's first-order beliefs: $T_i \mapsto \Delta(S_{-i} \times T_{-i}) \mapsto \Delta(S_{-i})$ (marginalizing)
2. i's second-order beliefs: $T_i \mapsto \Delta(S_{-i} \times T_{-i}) \mapsto \Delta(S^{-i} \times \times_{i \neq j} \Delta(S_{-j} \times T_{-j})) \mapsto \Delta(S_{-i} \times \times_{j \neq i} \Delta(S_{-j}))$ (marginalizing)

Literature

R. Myerson. *Harsanyi's Games with Incomplete Information*. Special 50th anniversary issue of *Management Science*, 2004.

M. Siniscalchi. *Epistemic Game Theory: Beliefs and Types*. New Palgrave Dictionary of Economics (forthcoming).

More on Types

- For any given set S of external states we can use a type space on S to provide consistent representations of the players' beliefs.

- Every state in a belief model or type space induces an infinite hierarchy of beliefs, but *not all consistent and coherent infinite hierarchies are in any finite model*. It is not obvious that even in an infinite model that all such hierarchies of beliefs can be represented.

more on this later

More on Types

- For any given set S of external states we can use a type space on S to provide consistent representations of the players' beliefs.
- Every state in a belief model or type space induces an infinite hierarchy of beliefs, but *not all consistent and coherent infinite hierarchies are in any finite model*. It is not obvious that even in an infinite model that all such hierarchies of beliefs can be represented.

more on this later

An Example

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

$\lambda_r(t_r)$

u_c	0	1/2
t_c	0	1/2
	L	R

$\lambda_r(u_r)$

u_c	1/2	0
t_c	0	1/2
	L	R

$\lambda_c(t_c)$

u_r	0	1/2
t_r	0	1/2
	U	D

$\lambda_c(u_c)$

u_r	1/2	0
t_r	0	1/2
	U	D

State: (D, t_r, R, t_c)

An Example

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

$\lambda_r(t_r)$

u_c	0	1/2
t_c	0	1/2
	L	R

$\lambda_r(u_r)$

u_c	1/2	0
t_c	0	1/2
	L	R

$\lambda_c(t_c)$

u_r	0	1/2
t_r	0	1/2
	U	D

$\lambda_c(u_c)$

u_r	1/2	0
t_r	0	1/2
	U	D

State: (D, t_r, R, t_c)

An Example

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

$\lambda_r(t_r)$

u_c	0	1/2
t_c	0	1/2
	L	R

$\lambda_r(u_r)$

u_c	1/2	0
t_c	0	1/2
	L	R

$\lambda_c(t_c)$

u_r	0	1/2
t_r	0	1/2
	U	D

$\lambda_c(u_c)$

u_r	1/2	0
t_r	0	1/2
	U	D

r is **correct** about c's strategy (similarly, for c).

An Example

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

$\lambda_r(t_r)$

u_c	0	1/2
t_c	0	1/2
	L	R

$\lambda_r(u_r)$

u_c	1/2	0
t_c	0	1/2
	L	R

$\lambda_c(t_c)$

u_r	0	1/2
t_r	0	1/2
	U	D

$\lambda_c(u_c)$

u_r	1/2	0
t_r	0	1/2
	U	D

r thinks it is possible c is wrong about her strategy

An Example

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

$\lambda_r(t_r)$

u_c	0	$1/2$
t_c	0	$1/2$
	L	R

$\lambda_r(u_r)$

u_c	$1/2$	0
t_c	0	$1/2$
	L	R

$\lambda_c(t_c)$

u_r	0	$1/2$
t_r	0	$1/2$
	U	D

$\lambda_c(u_c)$

u_r	$1/2$	0
t_r	0	$1/2$
	U	D

r is **rational**. (Similarly for c)

An Example

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

$\lambda_r(t_r)$

u_c	0	$1/2$
t_c	0	$1/2$
	L	R

$\lambda_r(u_r)$

u_c	$1/2$	0
t_c	0	$1/2$
	L	R

$\lambda_c(t_c)$

u_r	0	$1/2$
t_r	0	$1/2$
	U	D

$\lambda_c(u_c)$

u_r	$1/2$	0
t_r	0	$1/2$
	U	D

r thinks it is possible that c is **irrational**.

Expectation 1: Rationality and common belief of rationality

- What happens if all players are rational, believe that all players are rational, believe that all players believe that (...)?

- "Classical" assumption about game-theoretic analysis. See e.g. Myerson (1991).

Expectation 1: Rationality and common belief of rationality

- What happens if all players are rational, believe that all players are rational, believe that all players believe that (...)?

- "Classical" assumption about game-theoretic analysis. See e.g. Myerson (1991).

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- B is a bad strategy for Bob.
- It is *never* rational for him to choose B.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- B is a bad strategy for Bob.
- It is *never* rational for him to choose B.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- B is a bad strategy for Bob.
- It is *never* rational for him to choose B.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0$!

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0!$

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$EV_{t_B}(B) \geq EV_{t_B}(A)$$

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0!$

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$v_{Bob}(aB)\lambda_{Bob}(t_{Bob})(a) + v_{Bob}(bB)\lambda_{Bob}(t_{Bob})(b) \geq v_{Bob}(aA)\lambda_{Bob}(t_{Bob})(a) + v_{Bob}(bA)\lambda_{Bob}(t_{Bob})(b)$$

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0!$

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$1\lambda_{Bob}(t_{Bob})(a) + 0\lambda_{Bob}(t_{Bob})(b) \geq 2\lambda_{Bob}(t_{Bob})(a)$$

$$+1\lambda_{Bob}(t_{Bob})(b)$$

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0$!

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$\lambda_{Bob}(t_{Bob})(a) \geq 2\lambda_{Bob}(t_b)(a)$$
$$+\lambda_{Bob}(t_{Bob})(b)$$

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0!$

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0$!

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- A type t_B of Bob would be rational in choosing B iff:

$$0 \geq \lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b)$$

But then $\lambda_{Bob}(t_{Bob})(a) + \lambda_{Bob}(t_{Bob})(b) = 0$!

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- Bob never plays **B** at state (σ, t) if he is rational at that state.
- But then if Ann's type at that state believes that Bob is rational, that type must assign probability 1 to Bob playing **A**.
- Given this belief, **a** is her only rational strategy.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- Bob never plays **B** at state (σ, t) if he is rational at that state.
- But then if Ann's type at that state believes that Bob is rational, that type must assign probability 1 to Bob playing **A**.
- Given this belief, **a** is her only rational strategy.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- Bob never plays **B** at state (σ, t) if he is rational at that state.
- But then if Ann's type at that state believes that Bob is rational, that type must assign probability 1 to Bob playing **A**.
- Given this belief, **a** is her only rational strategy.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- Bob never plays **B** at state (σ, t) if he is rational at that state.
- But then if Ann's type at that state believes that Bob is rational, that type must assign probability 1 to Bob playing **A**.
- Given this belief, **a** is her only rational strategy.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- If Ann and Bob are rational, and Ann believes that Bob is rational at state (σ, t), then $\sigma = aA$.
- This strategy profile is the only one that survives *iterated elimination of strictly dominated strategies*.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- If Ann and Bob are rational, and Ann believes that Bob is rational at state (σ, t), then $\sigma = aA$.
- This strategy profile is the only one that survives *iterated elimination of strictly dominated strategies*.

Example

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

- If Ann and Bob are rational, and Ann believes that Bob is rational at state (σ, t), then $\sigma = \text{aA}$.
- This strategy profile is the only one that survives *iterated elimination of strictly dominated strategies*.

Strictly dominated strategies

Definition

A strategy s_i is *strictly dominated* by another strategy s'_i iff for all combinations of choices of the other players σ_{-i} :

$$v_i(s_i, \sigma_{-i}) < v_i(s'_i, \sigma_{-i})$$

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

	A	B
a	1, 2	0, 1
b	0, 1	1, 0

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

	A
a	1, 2
b	0, 1

Iterated elimination of strictly dominated strategies

1. Start with a game;
2. Eliminate all strictly dominated strategies;
3. Look at the reduced game;
4. Eliminate all strictly dominated strategies here;
5. Repeat 3 and 4 until you don't eliminate anything.

	A
a	1, 2

Common knowledge of rational and elimination of strictly dominated strategies

- If Ann and Bob are rational, and Ann believes that Bob is rational at state (σ, t), then $\sigma =$ aA.
- For this game we need rationality and only one level of higher-order information to conclude that aA will be played. But in the general case:

Theorem

For any state (σ, t) of a type structure for an arbitrary finite game $\mathbb{G}$, if all players are rational and it is common belief that all players are rational at (σ, t), then σ is a iteratively non-dominated strategy profile.

A. Brandenburger and E. Dekel. *Rationalizability and correlated equilibria. Econometrica*, 55:1391-1402, 1987.

Common knowledge of rational and elimination of strictly dominated strategies

- If Ann and Bob are rational, and Ann believes that Bob is rational at state (σ, t), then $\sigma = \text{aA}$.
- For this game we need rationality and only one level of higher-order information to conclude that aA will be played. But in the general case:

Theorem

For any state (σ, t) of a type structure for an arbitrary finite game $\mathbb{G}$, if all players are rational and it is common belief that all players are rational at (σ, t), then σ is a iteratively non-dominated strategy profile.

A. Brandenburger and E. Dekel. *Rationalizability and correlated equilibria. Econometrica*, 55:1391-1402, 1987.

Common knowledge of rational and elimination of strictly dominated strategies

- If Ann and Bob are rational, and Ann believes that Bob is rational at state (σ, t), then $\sigma = aA$.
- For this game we need rationality and only one level of higher-order information to conclude that aA will be played. But in the general case:

Theorem

For any state (σ, t) of a type structure for an arbitrary finite game $\mathbb{G}$, if all players are rational and it is common belief that all players are rational at (σ, t), then σ is a iteratively non-dominated strategy profile.

A. Brandenburger and E. Dekel. *Rationalizability and correlated equilibria*. *Econometrica*, 55:1391-1402, 1987.

Epistemic Characterizations of Solutions Concepts

If the players all satisfy some **epistemic condition** involving some form of **rationality** (eg., common knowledge of rationality) then the players will play according to some solution concept (eg., Nash equilibrium, iterated removal of strongly dominated strategies, . . .).

Two key assumptions about the rationality of players:

1. Common *knowledge* of *rationality* (i.e., common knowledge of choosing optimally)
2. Common prior

Epistemic Characterizations of Solutions Concepts

If the players all satisfy some **epistemic condition** involving some form of **rationality** (eg., common knowledge of rationality) then the players will play according to some solution concept (eg., Nash equilibrium, iterated removal of strongly dominated strategies, ...).

Two key assumptions about the rationality of players:

1. Common *knowledge* of *rationality* (i.e., common knowledge of choosing optimally)
2. Common prior

Removal of Strictly Dominated Strategies

We have seen that *common knowledge of rationality* implies that the players will follow the process of iteratively removing strictly dominated strategies.

1. Players should not choose strictly dominated strategies (*it is never rational*)
2. Assuming the above statement is common knowledge is equivalent to assuming the players iteratively remove strictly dominated strategies.

Removal of Strictly Dominated Strategies

We have seen that *common knowledge of rationality* implies that the players will follow the process of iteratively removing strictly dominated strategies.

1. Players should not choose strictly dominated strategies (*it is never rational*)
2. Assuming the above statement is common knowledge is equivalent to assuming the players iteratively remove strictly dominated strategies.

Admissibility

Can the same be proven for *admissibility*, i.e., avoidance of *weakly* dominated strategies?

Admissibility

Can the same be proven for *admissibility*, i.e., avoidance of *weakly* dominated strategies?

	L	R
T	$1, 1$	$0, 0$
M	$1, 1$	$2, 1$
B	$0, 0$	$2, 1$

Admissibility

Can the same be proven for *admissibility*, i.e., avoidance of *weakly* dominated strategies?

	L	R
T	$1, 1$	$0, 0$
M	$1, 1$	$2, 1$
B	$0, 0$	$2, 1$

Admissibility

Does assuming that it is commonly known that players play only admissible strategies lead to a process of iterated removal of weakly dominated strategies? no!

L. Samuelson. *Dominated Strategies and Common Knowledge*. Games and Economic Behavior (1992).

Admissibility

Does assuming that it is commonly known that players play only admissible strategies lead to a process of iterated removal of weakly dominated strategies? no!

L. Samuelson. *Dominated Strategies and Common Knowledge*. Games and Economic Behavior (1992).

Results

1. Removal of weakly dominated strategies and common knowledge of admissibility diverge.

2. There exist games in which assuming that admissibility is common knowledge does not provide players with sufficient information to determine which strategies should be eliminated on admissibility grounds.

3. There exists games in which assuming that admissibility is common knowledge yields a contradiction

L. Samuelson. *Dominated Strategies and Common Knowledge*. Games and Economic Behavior (1992).

Results

1. Removal of weakly dominated strategies and common knowledge of admissibility diverge.

2. There exist games in which assuming that admissibility is common knowledge does not provide players with sufficient information to determine which strategies should be eliminated on admissibility grounds.

3. There exists games in which assuming that admissibility is common knowledge yields a contradiction

L. Samuelson. *Dominated Strategies and Common Knowledge*. Games and Economic Behavior (1992).

Results

1. Removal of weakly dominated strategies and common knowledge of admissibility diverge.

2. There exist games in which assuming that admissibility is common knowledge does not provide players with sufficient information to determine which strategies should be eliminated on admissibility grounds.

3. There exists games in which assuming that admissibility is common knowledge yields a contradiction

L. Samuelson. *Dominated Strategies and Common Knowledge*. Games and Economic Behavior (1992).

An Issue

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

Suppose rationality incorporates *admissibility* (or *cautiousness*).

1. Both Row and Column should use a *full-support* probability measure
2. But if Row thinks that Column is **rational** then should she not assign probability 1 to L?

The condition that the players are rational *seems to conflict with the condition that the players think the other players are rational*

An Issue

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

Suppose rationality incorporates *admissibility* (or *cautiousness*).

1. Both Row and Column should use a *full-support* probability measure
2. But if Row thinks that Column is **rational** then should she not assign probability 1 to L?

The condition that the players are rational *seems to conflict with the condition that the players think the other players are rational*

An Issue

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

Suppose rationality incorporates *admissibility* (or *cautiousness*).

1. Both Row and Column should use a *full-support* probability measure
2. But if Row thinks that Column is **rational** then should she not assign probability 1 to L?

The condition that the players are rational *seems to conflict with the condition that the players think the other players are rational*

An Issue

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

Suppose rationality incorporates *admissibility* (or *cautiousness*).

1. Both Row and Column should use a *full-support* probability measure
2. But if Row thinks that Column is **rational** then should she not assign probability 1 to L?

The condition that the players are rational *seems to conflict with the condition that the players think the other players are rational*

An Issue

The argument for deletion of a weakly dominated strategy for player i is that he contemplates the possibility that every strategy combination of his rivals occurs with positive probability. However, this hypothesis clashes with the logic of iterated deletion, which assumes, precisely, that eliminated strategies are not expected to occur.

Mas-Colell, Whinston and Green. *Introduction to Microeconomics*. 1995.

Both Including and Excluding a Strategy

One solution is to assume that players consider some strategies *infinitely more likely than other strategies*.

Lexiographic Probability System: a sequence of probability distributions each infinitely more likely than the next.

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

Both Including and Excluding a Strategy

One solution is to assume that players consider some strategies *infinitely more likely than other strategies*.

Lexiographic Probability System: a sequence of probability distributions each infinitely more likely than the next.

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

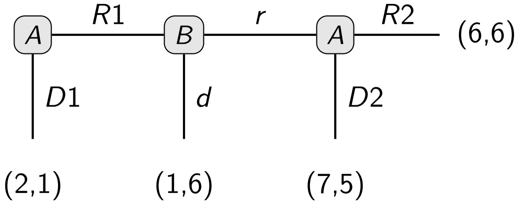
Both Including and Excluding a Strategy

One solution is to assume that players consider some strategies *infinitely more likely than other strategies*.

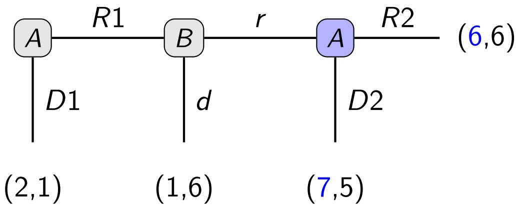
Lexiographic Probability System: a sequence of probability distributions each infinitely more likely than the next.

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

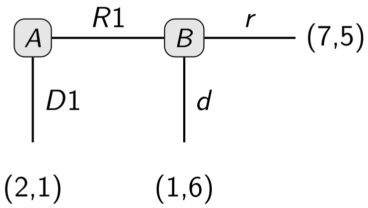
BI Puzzle



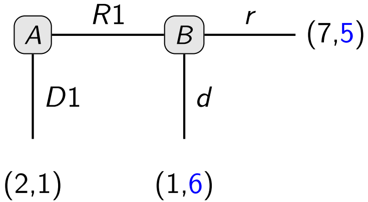
BI Puzzle



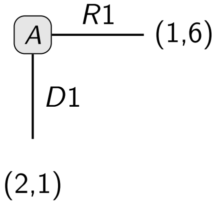
BI Puzzle



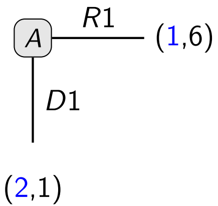
BI Puzzle



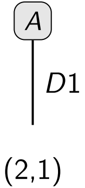
BI Puzzle



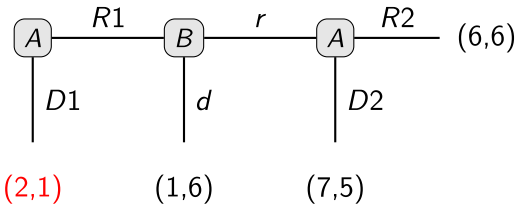
BI Puzzle



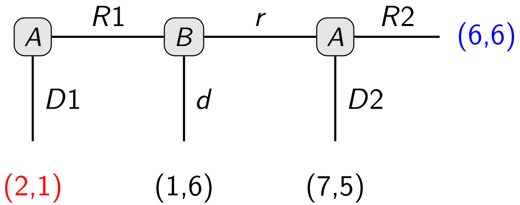
BI Puzzle



BI Puzzle

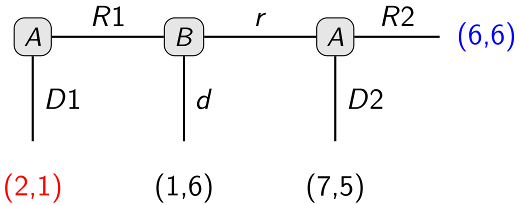


But what if...



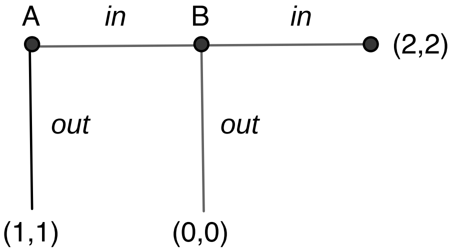
- Are the players *irrational*?
- What argument leads to the BI solution?

But what if...



- Are the players *irrational*?
- What argument leads to the BI solution?

A Problem: Probability Zero Events



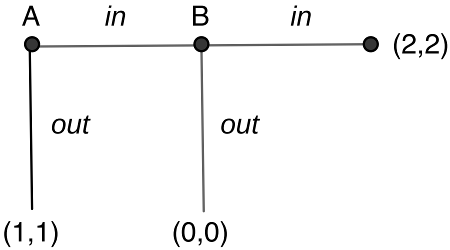
$\lambda_A(t_A)$

t_B	1	0
	out	*in*

$\lambda_B(t_B)$

t_A	1	0
	out	*in*

A Problem: Probability Zero Events



$\lambda_A(t_A)$

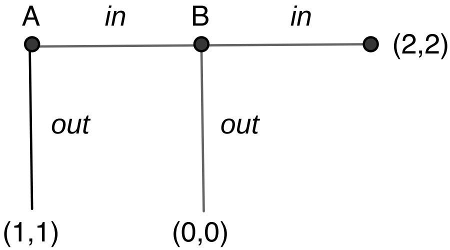
t_B	1	0
	out	*in*

$\lambda_B(t_B)$

t_A	1	0
	out	*in*

It can easily check that Ann and Bob or rational and commonly know each other are rational.

A Problem: Probability Zero Events



$\lambda_A(t_A)$

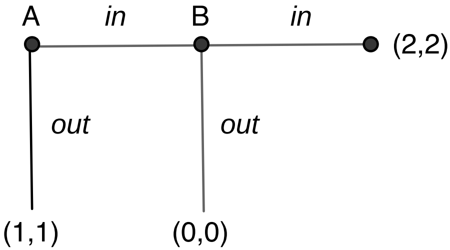
t_B	1	0
	out	*in*

$\lambda_B(t_B)$

t_A	1	0
	out	*in*

Bob plays Out because he is sure that Ann will also play out and so is indifferent between his moves.

A Problem: Probability Zero Events



$\lambda_A(t_A)$

t_B	1	0
	out	*in*

$\lambda_B(t_B)$

t_A	1	0
	out	*in*

BUT Ann knows that if she plays In then Bob will see this and so needs to think about how Bob will react.

Characterizing Backward Induction

Aumann's Theorem Common knowledge of substantive rationality implies the backward induction solution in games of perfect information.

Stalnaker's Theorem Common knowledge of substantive rationality does not imply the backward induction solution in games of perfect information.

substantive rationality: for nodes *n*, if the player were to reach node *n* then the players would be rational at *n*.

The *only* difference between Aumann and Stalnaker is how they interpret the above conditional.

J. Halpern. *Substantive Rationality and Backward Induction*. Games and Economic Behavior, 1998.

Characterizing Backward Induction

Aumann's Theorem Common knowledge of substantive rationality implies the backward induction solution in games of perfect information.

Stalnaker's Theorem Common knowledge of substantive rationality does not imply the backward induction solution in games of perfect information.

substantive rationality: for nodes *n*, if the player were to reach node *n* then the players would be rational at *n*.

The *only* difference between Aumann and Stalnaker is how they interpret the above conditional.

J. Halpern. *Substantive Rationality and Backward Induction*. Games and Economic Behavior, 1998.

Characterizing Backward Induction

Aumann's Theorem Common knowledge of substantive rationality implies the backward induction solution in games of perfect information.

Stalnaker's Theorem Common knowledge of substantive rationality does not imply the backward induction solution in games of perfect information.

substantive rationality: for nodes n, if the player were to reach node n then the players would be rational at n.

The *only* difference between Aumann and Stalnaker is how they interpret the above conditional.

J. Halpern. *Substantive Rationality and Backward Induction*. Games and Economic Behavior, 1998.

Characterizing Backward Induction

Aumann's Theorem Common knowledge of substantive rationality implies the backward induction solution in games of perfect information.

Stalnaker's Theorem Common knowledge of substantive rationality does not imply the backward induction solution in games of perfect information.

substantive rationality: for nodes n, if the player were to reach node n then the players would be rational at n.

The *only* difference between Aumann and Stalnaker is how they interpret the above conditional.

J. Halpern. *Substantive Rationality and Backward Induction*. Games and Economic Behavior, 1998.

Other Characterizations of BI

A. Perea. *Survey of Epistemic Characterizations of Backwards Induciton*. Interactive Logic, 2007.

Probability Zero Events and Type Spaces

Two main approaches have been put forward to deal with reasoning about probability zero events:

Lexiographic Probability Systems: a sequence of probability distributions each infinitely more likely than the next.

Conditional Probability Systems: for events E, p_E is a probability distribution (even if $p_\Omega(E) = 0$).

LPS used in the characterization of IA on strategic games

CPS is used for BI on extensive games

Both analyses require *large* type spaces.

J. Halpern. *Lexicographic probability, conditional probability and non-standard probability*. .

Types of Irrationality?

1. A player is irrational if he does not optimize given its current beliefs
2. A player is irrational if, although he optimizes, he does not consider *all possibilities*

A. Brandenburger, A. Friedenberg, H. J. Keisler. *Admissibility in Games*. Econometrica (2008).

Types of Irrationality?

1. A player is irrational if he does not optimize given its current beliefs
2. A player is irrational if, although he optimizes, he does not consider *all possibilities*

A. Brandenburger, A. Friedenberg, H. J. Keisler. *Admissibility in Games*. Econometrica (2008).

The General Question

Does such a space of all possible (interactive) beliefs exist?

A Question

- For any given set S of external states we can use a Bayesian model or a type space on S to provide consistent representations of the players' beliefs.

- Every state in a belief model or type space induces an infinite hierarchy of beliefs, but *not all consistent infinite hierarchies are in any finite model*. It is not obvious that even in an infinite model that all consistent hierarchies of beliefs can be represented.

- Which type space is the "correct" one to work with?

A Question

- For any given set S of external states we can use a Bayesian model or a type space on S to provide consistent representations of the players' beliefs.
- Every state in a belief model or type space induces an infinite hierarchy of beliefs, but *not all consistent infinite hierarchies are in any finite model*. It is not obvious that even in an infinite model that all consistent hierarchies of beliefs can be represented.
- Which type space is the "correct" one to work with?

A Question

- For any given set S of external states we can use a Bayesian model or a type space on S to provide consistent representations of the players' beliefs.

- Every state in a belief model or type space induces an infinite hierarchy of beliefs, but *not all consistent infinite hierarchies are in any finite model*. It is not obvious that even in an infinite model that all consistent hierarchies of beliefs can be represented.

- Which type space is the "correct" one to work with?

Is there a universal type space?

A **universal type space** is a types space to which every type space (on the same space of states of nature and same set of agents) can be mapped, preferably in a unique way, by a map that preserves the structure of the type space.

If such a space exists, then the any analysis of a game could be carried out in this space without the risk of missing any "relevant" states of affairs.

Yes, if ...

The existence of a universal types space depends on the topological and/or measure theoretic assumptions being made about the underlying state space S.

First shown by Mertens and Zamir (1985)

The problem is to define the set of all infinite hierarchies of beliefs satsifying the same consistency properties (coherency and common knowledge of coherency) as that of hierarchies obtained at some state in a type space.

Kolomogorov Extension Theorem

Yes, if ...

The existence of a universal types space depends on the topological and/or measure theoretic assumptions being made about the underlying state space S.

First shown by Mertens and Zamir (1985)

The problem is to define the set of all infinite hierarchies of beliefs satsifying the same consistency properties (coherency and common knowledge of coherency) as that of hierarchies obtained at some state in a type space.

Kolomogorov Extension Theorem

Why do we care?

It turns out that finding the connection between rationality, what agents think about the situation and what actually happens depends on the existence of a "rich enough" space of types, i.e., a universal type space.

> It is not enough [...] that Ann should consider each of Bob's strategies possible. Rather, she considers possible both every strategy that Bob might play and every type that Bob might be. (Likewise, Bob considers possible both every strategy that Ann might play and every type that Ann might be.)

Brandenburger, Friedenburg and Keisler. *Admissibility in Games*. 2004.

Brandenburger and Keisler. *Epistemic Conditions for Iterated Admissibility*. Proceedings of TARK 2001.

Why do we care?

It turns out that finding the connection between rationality, what agents think about the situation and what actually happens depends on the existence of a "rich enough" space of types, i.e., a universal type space.

> It is not enough [...] that Ann should consider each of Bob's strategies possible. Rather, she considers possible both every strategy that Bob might play and every type that Bob might be. (Likewise, Bob considers possible both every strategy that Ann might play and every type that Ann might be.)

Brandenburger, Friedenburg and Keisler. *Admissibility in Games*. 2004.

Brandenburger and Keisler. *Epistemic Conditions for Iterated Admissibility*. Proceedings of TARK 2001.

Canonical, Complete and Terminal Models

1. **Canonical models**: Start with a space of underlying uncertainty, players form beliefs over this space, believes over this space and the space of 0-th order beliefs, and so on inductively. The question is, does this process end?
2. **Complete models**: The "two-way subjectivity" models described later.
3. **Terminal models**: Given a category **C** of models of beliefs, call a model $\mathcal{M}$ in **C** terminal if for any other model $\mathcal{N}$ in **C**, there is a unique belief preserving morphism from $\mathcal{M}$ to $\mathcal{N}$.

Overview of the Literature

- Existence proofs (under various topological assumptions): [Armbruster and Böge, 1979], [Mertens and Zamir, 1985], [Brandenburger and Dekel, 1993], [Heifetz, 1993], [Heifetz and Samet, 1998], [Battigalli and Siniscalchi, 1999], [Meier, 2002], [Salonen, 2003]

- Impossibility Result: [Brandenburger and Kesiler, 2004], [Meier, 2005]

Some Literature

A. Brandenburger and E. Dekel. *Hierarchies of Beliefs and Common Knowledge*. Journal of Economic Theory (1993).

A. Heifetz and D. Samet. *Knoweldge Spaces with Arbitrarily High Rank*. Games and Economic Behavior (1998).

L. Moss and I. Viglizzo. *Harsanyi type spaces and final coalgebras constructed from satisfied theories*. EN in Theoretical Computer Science (2004).

A. Friendenberg. *When do type structures contain all hierarchies of beliefs?*. working paper (2007).

Some Literature

A. Brandenburger and E. Dekel. *Hierarchies of Beliefs and Common Knowledge*. Journal of Economic Theory (1993).

A. Heifetz and D. Samet. *Knoweldge Spaces with Arbitrarily High Rank*. Games and Economic Behavior (1998).

L. Moss and I. Viglizzo. *Harsanyi type spaces and final coalgebras constructed from satisfied theories*. EN in Theoretical Computer Science (2004).

A. Friendenberg. *When do type structures contain all hierarchies of beliefs?*. working paper (2007).

Some Literature

A. Brandenburger and E. Dekel. *Hierarchies of Beliefs and Common Knowledge*. Journal of Economic Theory (1993).

A. Heifetz and D. Samet. *Knoweldge Spaces with Arbitrarily High Rank*. Games and Economic Behavior (1998).

L. Moss and I. Viglizzo. *Harsanyi type spaces and final coalgebras constructed from satisfied theories*. EN in Theoretical Computer Science (2004).

A. Friendenberg. *When do type structures contain all hierarchies of beliefs?*. working paper (2007).

Some Literature

A. Brandenburger and E. Dekel. *Hierarchies of Beliefs and Common Knowledge*. Journal of Economic Theory (1993).

A. Heifetz and D. Samet. *Knoweldge Spaces with Arbitrarily High Rank*. Games and Economic Behavior (1998).

L. Moss and I. Viglizzo. *Harsanyi type spaces and final coalgebras constructed from satisfied theories*. EN in Theoretical Computer Science (2004).

A. Friendenberg. *When do type structures contain all hierarchies of beliefs?*. working paper (2007).

Brandenburger and Dekel

Assumption: Assume there are only two agents: i, j. Let the state space S be a Polish space (complete separable metric). For any metric space X assume that $\Delta(X)$ is endowed with the weak topology.

The proof proceeds as follows

1. Inductively construct the set of all possible types. Formally, types are infinite sequences of probability measures.
2. Define a notion of *coherency* such that if an individual's type is assumed to be coherent then it induces a belief over the types of the other idividuals.
3. If common knowledge (in the sense of assigning probability 1) of coherency is assumed, then the set of beliefs is closed.

Brandenburger and Dekel

Assumption: Assume there are only two agents: i, j. Let the state space S be a Polish space (complete separable metric). For any metric space X assume that $\Delta(X)$ is endowed with the weak topology.

The proof proceeds as follows

1. Inductively construct the set of all possible types. Formally, types are infinite sequences of probability measures.
2. Define a notion of *coherency* such that if an individual's type is assumed to be coherent then it induces a belief over the types of the other idividuals.
3. If common knowledge (in the sense of assigning probability 1) of coherency is assumed, then the set of beliefs is closed.

Brandenburger and Dekel

Assumption: Assume there are only two agents: i, j. Let the state space S be a Polish space (complete separable metric). For any metric space X assume that $\Delta(X)$ is endowed with the weak topology.

The proof proceeds as follows

1. Inductively construct the set of all possible types. Formally, types are infinite sequences of probability measures.

2. Define a notion of *coherency* such that if an individual's type is assumed to be coherent then it induces a belief over the types of the other idividuals.

3. If common knowledge (in the sense of assigning probability 1) of coherency is assumed, then the set of beliefs is closed.

Step 1.

$$
\begin{aligned}
X_0 &= S \\
X_1 &= X_0 \times \Delta(X_0) \\
&\vdots \\
X_n &= X_{n-1} \times \Delta(X_{n-1}) \\
&\vdots
\end{aligned}
$$

A *type* t^i of i is an infinite sequence $t^i = (\delta^i_1, \delta^i_2, \ldots) \in \Pi_{n=0}^{\infty} \Delta(X_n)$

Let $T_0 = \Pi_{n=0}^{\infty} \Delta(X_n)$.

Step 2.

Coherent: A type $t = (\delta_1, \delta_2, \ldots) \in T_0$ is *coherent* if for every $n \geq 2$, $marg_{X_{n-2}} \delta_n = \delta_{n-1}$.

Coherency simply says that different levels of beliefs of an individual do not contradict one another. Let T_1 be the set of all coherent types.

Proposition There is a homeomorphism $f : T_1 \to \Delta(S \times T_0)$.

This is essentially *Kolmogorov's Existence Theorem*.

Note that the marginal probability assigned by $f(\delta_1, \delta_2, \ldots)$ to a given event in X_{n-1} is equal to the probability that δ_n assigns to that same event.

Step 3.

We now impose "common knowledge" of coherency:

For $k \geq 2$ define

$$T_k = \{t \in T_1 : f(t)(S \times T_{k-1}) = 1\}$$

Let $T = \cap_{k=1}^{\infty} T_k$

This set T is the set we are looking for: the universal type space.

Proposition There is a homeomorphism $g : T \rightarrow \Delta(S \times T)$

Aumann's probability logic

Let Φ be a set of proposition letters and Agt a set of agents. Let $\mathcal{L}_A$ be the language with the following formulas:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid L^i_r(\varphi)$$

where $p \in \Phi$, $i \in Agt$, and $r \in \mathbb{Q} \cap [0, 1]$.

- $L^i_r(\varphi)$ is read "the probability of φ is at least r".
- $M^i_r(\varphi) \equiv L^i_{1-r}(\neg\varphi)$ is read "the probability of φ is at most r".

Harsanyi types

Let Δ be a function that maps each measurable space $(X, \mathcal{A})$ to the measurable space $(Y, \mathcal{B})$, where

- Y is the set of probability measures on $(X, \mathcal{A})$
- $\mathcal{B}$ is the σ-algebra generated by $\{\beta^p(A) : p \in \mathbb{Q} \cap [0,1], A \in \mathcal{A}\}$, where

$$\beta^p(A) = \{\mu : \mu(A) \geq p\}.$$

Let I_0 be a set of players, and $I = I_0 \cup \{0\}$.
Let U_i map a family $(X_j)_{j \in I}$ of measurable spaces to the product $\prod\{X_j : j \in I_0, i \neq j\}$.

Definition

A Type space is a family $X = (X_j)_{j \in I}$ of measurable spaces together with a family $(f_i : X_i \longrightarrow \Delta U_i(X))_{i \in I}$ of functions.

Semantics

Let $T = (X, f)$ be a type space, with $X = (X_i)_{i \in I}$ and $f = (f_i)_{i \in I}$. Let $S = (S_i)_{i \in I}$ be the family of sample spaces. Augment T with a function $\| \cdot \|$ mapping a set Φ of proposition letters to $\mathcal{P}(S)$. Define a function $[\![\cdot]\!]$ from formulas to $\mathcal{P}(S)$, such that

$$
\begin{array}{lcl}
[\![\top]\!] & = & X \\
[\![p]\!] & = & \|p\| \\
[\![\neg\varphi]\!] & = & X - [\![\varphi]\!] \\
[\![\varphi \wedge \psi]\!] & = & [\![\varphi]\!] \cap [\![\psi]\!] \\
[\![L^i_r \varphi]\!] & = & (\pi_i^{-1} \circ f_i^{-1} \circ \beta_i^r \circ \pi_{-\{i,0\}})([\![\varphi]\!])
\end{array}
$$

One agent version

Set Φ of proposition letters.
Models: $M = (X, \mathcal{A}, f, \nu)$, where

- $(X,)$ is a measurable space
- $f : X \longrightarrow \Delta(X, \mathcal{A}$ is measurable
- $\nu : \Phi \longrightarrow \mathcal{P}(X)$

Then

$$[\![L_r\varphi]\!] = (f^{-1} \circ \beta^r)([\![\varphi]\!])$$

Proof system

- All propositional tautologies
- $L_0(\varphi)$, for all formulas φ
- $L_r(\top)$, for all $r \in \mathbb{Q} \cap [0,1]$
- $L_r\varphi \longrightarrow \neg L_s \neg\varphi$, for $r + s > 1$
- $L_r(\varphi \wedge \psi) \wedge L_s(\varphi \wedge \neg\psi) \longrightarrow L_{r+s}(\varphi)$, for $r + s \leq 1$
- $\neg L_r(\varphi \wedge \psi) \wedge \neg L_s(\varphi \wedge \neg\psi) \longrightarrow \neg L_{r+s}(\varphi)$, for $r + s \leq 1$
- If $\vdash \varphi \leftrightarrow \psi$, then $\vdash L_r\varphi \leftrightarrow L_r\psi$
- If $\vdash \gamma \longrightarrow L_s\varphi$ for all $s < r$, then $\vdash \gamma \longrightarrow L_r\varphi$
- If $\vdash \varphi$ and $\vdash \varphi \longrightarrow \psi$, then $\vdash \psi$.

This system is sound and weakly complete with respect to the one agent semantics.