The unknown Stochastic Game

Mario Bravo and Miquel Oliu-Barton

USACH

Université Paris Dauphine

CMM Santiago de Chile

March 16, 2017

The Sphynx

Tell me, stranger, what creature walks
with four legs in the morning
with two legs at noon
and with three legs at night ?

The Sphynx

Tell me, stranger, what creature walks
with four legs in the morning
with two legs at noon
and with three legs at night ?

Oedipus

A man !

The Sphynx

Tell me, game theorist, how do you play
a (long) repeated game
in the morning
at noon
at night ?

The Sphynx

Tell me, game theorist, how do you play
a (long) repeated game
in the morning
at noon
at night ?

The game theorist

Oh, good question, Sphynx !

Lots of people are trying to figure it out...

The Sphynx

Tell me, game theorist, how do you play
a (long) repeated game
in the morning
at noon
at night ?

The game theorist

Oh, good question, Sphynx !

Lots of people are trying to figure it out...

We don't play the same way at different times of the day !

Let me give you a brief answer first

A brief answer to the Sphynx

- **In the morning**: the game emerges from the darkness...
  - Bayesian games
  - Games with incomplete information

  Some data (parameters) are imperfectly known; players have beliefs on them. The game theorist deals with private information

A brief answer to the Sphynx

- **In the morning**: the game emerges from the darkness...
  - Bayesian games
  - Games with incomplete information

  Some data (parameters) are imperfectly known; players have beliefs on them. The game theorist deals with private information

- **At noon**: the game is perfectly visible to all players
  - N-player games
  - Stochastic games

  Here, the game theorist deals with recursive formulae

A brief answer to the Sphynx

- **In the morning**: the game emerges from the darkness...
  - Bayesian games
  - Games with incomplete information

  Some data (parameters) are imperfectly known; players have beliefs on them. The game theorist deals with private information

- **At noon**: the game is perfectly visible to all players
  - N-player games
  - Stochastic games

  Here, the game theorist deals with recursive formulae

- **At night**: the game is totally unknown...
  - the unknown game
  - the unknown stochastic game

  How many players are there? How many actions, what preferences do they have ? The game theorist deals with no-regret algorithms

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - ...
  - ...
  - ...

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - ...
  - ...
  - ...
- How many players ?
  - ...
  - ...
  - ...

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - ...
  - ...
  - ...
- How many players ?
  - ...
  - ...
  - ...
- What about their preferences ?
  - ...
  - ...
  - ...
  - ...

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - Cooperative games
  - Evolutionary games
  - Non-cooperative games

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - Cooperative games
  - Evolutionary games
  - Non-cooperative games
- How many players ?
  - 1 player
  - N players
  - Infinitely many

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - Cooperative games
  - Evolutionary games
  - Non-cooperative games
- How many players ?
  - 1 player
  - N players
  - Infinitely many
- What about their preferences ?
  - Identical
  - Zero-sum
  - Anything

General context

Game = interdependent strategic interaction between players

- What sort of interaction ?
  - Cooperative games
  - **Non-cooperative games**
  - Evolutionary games
- How many players ?
  - 1 player
  - N **players**
  - Infinitely many
- What about their preferences ?
  - Identical
  - **Zero-sum**
  - **Anything**

Non-cooperative N-player games

A non-cooperative N player game is described by a triplet (N, A, g), where

- N is a set of players
- $A = (A^i)_{i \in N}$ is a set of actions
- $g = (g^i)_{i \in N}$ is the payoff function, $g : A \to \mathbb{R}^N$

Non-cooperative N-player games

A non-cooperative N player game is described by a triplet (N, A, g), where

- N is a set of players
- $A = (A^i)_{i \in N}$ is a set of actions
- $g = (g^i)_{i \in N}$ is the payoff function, $g : A \to \mathbb{R}^N$

An equilibrium is an action profile $a \in A$ such that

$$g^i(a^i, a^{-i}) \geq g^i(b^i, a^{-i}), \quad \forall b^i \in A^i, \ \forall i \in N$$

If a is an equilibrium, $g(a) \in \mathbb{R}^N$ is an equilibrium payoff

We are interested in the 3 following questions:

Non-cooperative N-player games

A non-cooperative N player game is described by a triplet (N, A, g), where

- N is a set of players
- $A = (A^i)_{i \in N}$ is a set of actions
- $g = (g^i)_{i \in N}$ is the payoff function, $g : A \to \mathbb{R}^N$

An equilibrium is an action profile $a \in A$ such that

$$g^i(a^i, a^{-i}) \geq g^i(b^i, a^{-i}), \quad \forall b^i \in A^i, \ \forall i \in N$$

If a is an equilibrium, $g(a) \in \mathbb{R}^N$ is an equilibrium payoff

We are interested in the 3 following questions:

(a) The set of equilibria

(b) The set of equilibrium payoffs

Non-cooperative N-player games

A non-cooperative N player game is described by a triplet (N, A, g), where

- N is a set of players
- $A = (A^i)_{i \in N}$ is a set of actions
- $g = (g^i)_{i \in N}$ is the payoff function, $g : A \to \mathbb{R}^N$

An equilibrium is an action profile $a \in A$ such that

$$g^i(a^i, a^{-i}) \geq g^i(b^i, a^{-i}), \quad \forall b^i \in A^i, \ \forall i \in N$$

If a is an equilibrium, $g(a) \in \mathbb{R}^N$ is an equilibrium payoff

We are interested in the 3 following questions:

(a) The set of equilibria

(b) The set of equilibrium payoffs

(c) The security levels $v^i = \sup_{a^i \in A^i} \inf_{a^{-i} \in A^{-i}} g^i(a^i, a^{-i})$, $i \in N$

Two-player zero-sum games

A 2-player zero-sum game is described by a triplet (N, A, g), where

- $N = \{1, 2\}$
- $A = (A^1, A^2)$
- $g = (g^1, g^2)$ with $g^1 + g^2 = 0$

If a is an equilibrium, $g(a) = (v, -v) \in \mathbb{R}^2$, and v is called the value
An optimal strategy is an action $a^i \in A^i$ such that

$$g^i(a^i, a^{-i}) \geq v, \quad \forall a^{-i} \in A^{-i}$$

Two-player zero-sum games

A 2-player zero-sum game is described by a triplet (N, A, g), where

- $N = \{1, 2\}$
- $A = (A^1, A^2)$
- $g = (g^1, g^2)$ with $g^1 + g^2 = 0$

If a is an equilibrium, $g(a) = (v, -v) \in \mathbb{R}^2$, and v is called the value
An optimal strategy is an action $a^i \in A^i$ such that

$$g^i(a^i, a^{-i}) \geq v, \quad \forall a^{-i} \in A^{-i}$$

Transposition to zero-sum games:

(a) Equilibria = couples of optimal strategies

(b) A unique equilibrium payoff = the value

(c) Security levels = the value

Battle of sexes (at noon)

- The game (N, A, g) is commonly known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

Battle of sexes (at noon)

- The game (N, A, g) is commonly known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

- The set of equilibria is $\{(R, R), (T, T), (\frac{1}{4}, \frac{3}{4}; \frac{3}{4}, \frac{1}{4})\}$
- The set of equilibrim payoffs is $\{(3, 2), (2, 3), (\frac{3}{2}, \frac{3}{2})\}$
- The security level for both players is $\frac{3}{2}$, the value of the game:

3	1
0	2

Battle of sexes (at noon)

- The game (N, A, g) is commonly known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

- The set of equilibria is $\{(R, R), (T, T), (\frac{1}{4}, \frac{3}{4}; \frac{3}{4}, \frac{1}{4})\}$
- The set of equilibrim payoffs is $\{(3, 2), (2, 3), (\frac{3}{2}, \frac{3}{2})\}$
- The security level for both players is $\frac{3}{2}$, the value of the game:

3	1
0	2

What if the game is played over and over ?

Battle of sexes (at noon)

- The game (N, A, g) is commonly known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

- The set of equilibria is $\{(R, R), (T, T), (\frac{1}{4}, \frac{3}{4}; \frac{3}{4}, \frac{1}{4})\}$
- The set of equilibrim payoffs is $\{(3, 2), (2, 3), (\frac{3}{2}, \frac{3}{2})\}$
- The security level for both players is $\frac{3}{2}$, the value of the game:

3	1
0	2

What if the game is played over and over ? **Folk theorems...**

Battle of sexes (in the morning)

- Several possible games (N, A, g^k), $k \in K$, which are known
- A prior belief $p \in \Delta(K)$ which is known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

p

	R	T
R	$3, 0$	$1, 3$
T	$0, 2$	$2, 0$

$1 - p$

Battle of sexes (in the morning)

- Several possible games (N, A, g^k), $k \in K$, which are known
- A prior belief $p \in \Delta(K)$ which is known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

p

	R	T
R	$3, 0$	$1, 3$
T	$0, 2$	$2, 0$

$1 - p$

What if the (true) game is played over and over ?

Battle of sexes (in the morning)

- Several possible games (N, A, g^k), $k \in K$, which are known
- A prior belief $p \in \Delta(K)$ which is known

	R	T
R	$3, 2$	$1, 1$
T	$0, 0$	$2, 3$

p

	R	T
R	$3, 0$	$1, 3$
T	$0, 2$	$2, 0$

$1 - p$

What if the (true) game is played over and over ?

"Cav u" theorem in the zero-su case, and many many extensions...

Battle of sexes (night)

- The game (N, A, g) is unknown

	?	?	...
R	?	?	
T	?	?	

Battle of sexes (night)

- The game (N, A, g) is unknown

	?	?	...
R	?	?	
T	?	?	

What if the game is repeated ?

The player plays $a_1, a_2, a_3, \ldots$ in $\{R, T\}$
He observes $g_1, g_2, g_3, \ldots$ where $g_t = g(a_t, b_t)$

Can player 1 ensure that his average payoff is at least $v^1 = \frac{3}{2}$?

Battle of sexes (night)

- The game (N, A, g) is unknown

	?	?	...
R	?	?	
T	?	?	

What if the game is repeated ?

The player plays $a_1, a_2, a_3, \ldots$ in $\{R, T\}$
He observes $g_1, g_2, g_3, \ldots$ where $g_t = g(a_t, b_t)$

Can player 1 ensure that his average payoff is at least $v^1 = \frac{3}{2}$?

Yes ! [Auer, Cesa-Bianchi, Freund, Schapire 1995]
No regret strategy (p.e. exponential weight algorithm) ensures the value

Let us move one step further

- So far, the game is fixed once and for all
- What if the game evolves during the play ?
- A simple model of dynamic game was proposed by [Shapley 1953]
- Can we still play it well under different information setups ?

The simplest dynamic model: stochastic games

Introduced by **Shapley 53**, stochastic games are described by a 5-tuple $\Gamma = (N, S, A, g, q)$ where

- N is a set of players
- S is a set of states
- $A = (A^i)_{i \in N}$ is a set of actions
- $g = (g^i)_{i \in N}$ is a stage payoff function, $g : S \times A \to \mathbb{R}^N$
- $q : S \times A \to \Delta(S)$ is a transition function

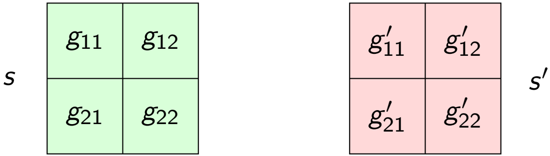
Outline of the game: at stage $m \geq 1$, knowing the current state s_m

- The players choose an action $a_m \in A$
- A stage-payoff $g_m := g(s_m, a_m)$ is produced
- A new state s_{m+1} is chosen according to $q(\cdot | s_m, a_m)$

From one-shot to stochastic games

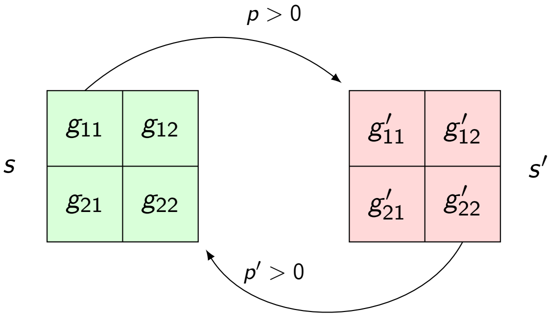
g_{11}	g_{12}
g_{21}	g_{22}

From one-shot to stochastic games



From one-shot to stochastic games

A stochastic game with two states and two actions



Dynamics and information

- Strategic interaction $\to$ flow of information + changes of state
  The state of the world evolves and the players notice it

- The players gather two types of information
  - Information about the game (the data)
  - Observation of the past play (actions and states)

- The times of play can be
  - Discrete: players interact at stages $m = 1, 2, 3, \ldots$.
  - Continuous (recent model from Neyman 2013)

Some comments and questions

Comments

- Stochastic games include several important classes of games (super-games, Markov decision processes, zero-sum stochastic games, quitting games, absorbing games and irreducible games,...)

Some comments and questions

Comments

- Stochastic games include several important classes of games (super-games, Markov decision processes, zero-sum stochastic games, quitting games, absorbing games and irreducible games,...)
- Trade-off between current and future reward

Some comments and questions

Comments

- Stochastic games include several important classes of games (super-games, Markov decision processes, zero-sum stochastic games, quitting games, absorbing games and irreducible games,...)
- Trade-off between current and future reward
- No loss of generality in the stationarity assumptions

Some comments and questions

Comments

- Stochastic games include several important classes of games (super-games, Markov decision processes, zero-sum stochastic games, quitting games, absorbing games and irreducible games,...)
- Trade-off between current and future reward
- No loss of generality in the stationarity assumptions

Questions

- How do players evaluate the stream of payoffs $(g^i(s_m, a_m))_{m\geq 1}$?

Some comments and questions

Comments

- Stochastic games include several important classes of games (super-games, Markov decision processes, zero-sum stochastic games, quitting games, absorbing games and irreducible games,...)
- Trade-off between current and future reward
- No loss of generality in the stationarity assumptions

Questions

- How do players evaluate the stream of payoffs $(g^i(s_m, a_m))_{m\geq 1}$?
- What information do the players have during the game ?

Some comments and questions

Comments

- Stochastic games include several important classes of games (super-games, Markov decision processes, zero-sum stochastic games, quitting games, absorbing games and irreducible games,...)
- Trade-off between current and future reward
- No loss of generality in the stationarity assumptions

Questions

- How do players evaluate the stream of payoffs $(g^i(s_m, a_m))_{m \geq 1}$?
- What information do the players have during the game ?
- What information do the players have about the game ?

The two extreme positions

- **Standard model**
  - The game (N, S, A, g, q) is commonly known
  - Perfect monitoring: the past play is observed

The two extreme positions

- **Standard model**
  - The game (N, S, A, g, q) is commonly known
  - Perfect monitoring: the past play is observed

- **Minimal information model**
  - The game (N, S, A, g, q) is unknown, player i knows only A^i
  - No prior belief on possible games
  - During the game, player i observes s_m and $g^i(s_m, a_m)$, $m = 1, 2, \ldots$

Applications

- Capital accumulation (e.g. fishery)
  - N players own a ressource (or a productive asset)
  - At each period they decide the amount of the ressource to consume
  - Literature: Levhari-Mirman 80, Dutta-Sundaram 93, Nowak 03...
- Taxation
  - A government sets a tax rate at every period
  - Each citizen decides how much to consume or save
  - Literature: Chari-Kehoe 90, Phelan-Stacchetti 01
- Others: Communication networks, queues,...

Questions

- How do players evaluate the stream of payoffs $(g^i(s_m, a_m))_{m \geq 1}$?
- What information do the players have during the game ?
- What information do the players have about the game ?

Answers

Questions

- How do players evaluate the stream of payoffs $(g^i(s_m, a_m))_{m \geq 1}$?
- What information do the players have during the game ?
- What information do the players have about the game ?

Answers

- It depends on the "horizon": finite, discounted and undiscounted
- Perfect monitoring versus imperfect monitoring
- Full information versus minimal information

Evaluation of the payoff

- A play (or history h) is a sequence of states and actions $(s_m, a_m)_{m \geq 1}$
- To each play corresponds a sequence of stage payoffs $(g(s_m, a_m))_m$

Discounted game (evaluation θ)

- Players evaluate their payoffs according to positive decreasing weights $g^i(h) = \sum_{m \geq 1} \theta^i_m g^i(s_m, a_m)$, where $\theta^i_m \geq 0$ and $\sum_{m \geq 1} \theta^i_m = 1$

Evaluation of the payoff

- A play (or history h) is a sequence of states and actions $(s_m, a_m)_{m\geq 1}$
- To each play corresponds a sequence of stage payoffs $(g(s_m, a_m))_m$

Discounted game (evaluation θ)

- Players evaluate their payoffs according to positive decreasing weights $g^i(h) = \sum_{m\geq 1} \theta^i_m g^i(s_m, a_m)$, where $\theta^i_m \geq 0$ and $\sum_{m\geq 1} \theta^i_m = 1$
- Two important cases: the n-stage game and the λ-discounted game

$$\theta^i_m = \frac{1}{n}\mathbb{1}_{\{m\leq n\}} \quad \text{and} \quad \theta^i_m = \lambda(1-\lambda)^{m-1}$$

Undiscounted game

- The players consider the $\liminf_{n\to+\infty} \frac{1}{n}\sum_{m=1}^{n} g^i(s_m, a_m)$

Strategies and Equilibria

- As usual, a strategy σ^i is a map from past history (or current information) to mixed actions
- A stationary strategy maps states into mixed actions, $\sigma^i : S \to \Delta(A^i)$
- A strategy profile $\sigma = (\sigma^i)_i$ induces a unique probability distribution over histories. Players maximize:

$$\gamma^i(s_1, \sigma) = \mathbb{E}_{s_1,\sigma,q}[\text{ discounted or undiscounted payoff }]$$

- A strategy profile σ is an equilibrium if

$$\gamma^i(s_1, \sigma^i, \sigma^{-i}) \geq \gamma^i(s_1, \rho^i, \sigma^{-i}), \quad \forall \tau^i \in \Sigma^i, \ \forall i \in N$$

- An equilibrium σ is stationary if σ^i is stationary for all i

Horizon and Equilibria

- **Fixed duration** (fixed evaluation)
  (a) ...
  (b) ...

- **Asymptotic approach** (evaluation tends to 0)
  (a) ...
  (b) ...

- **Uniform approach** (evaluation is "sufficiently close to 0")
  (a) ...
  (b) ...

Horizon and Equilibria

- **Fixed duration** (fixed evaluation)
  (a) Set of equilibria E_θ
  (b) Set of equilibrium payoffs P_θ

Horizon and Equilibria

- **Fixed duration** (fixed evaluation)
  - (a) Set of equilibria E_θ
  - (b) Set of equilibrium payoffs P_θ

- **Asymptotic approach** (evaluation tends to 0)
  - (a) Convergence of the set of equilibria $\lim_{\theta \to 0} E_\theta$
  - (b) Set of limit equilibrium payoffs $\lim_{\theta \to 0} P_\theta$

Horizon and Equilibria

- **Fixed duration** (fixed evaluation)
  - (a) Set of equilibria E_θ
  - (b) Set of equilibrium payoffs P_θ
- **Asymptotic approach** (evaluation tends to 0)
  - (a) Convergence of the set of equilibria $\lim_{\theta\to 0} E_\theta$
  - (b) Set of limit equilibrium payoffs $\lim_{\theta\to 0} P_\theta$
- **Uniform approach** (evaluation is "sufficiently close to 0")
  - (a) Existence of uniform ε-equilibria
  - (b) Characterization of the equilibrium payoffs

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series
- **Mertens and Neyman 81**: existence of the uniform value v_∞

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series
- **Mertens and Neyman 81**: existence of the uniform value v_∞
- **Bravo and O.-B 2017**: solve the unknown stochastic game

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series
- **Mertens and Neyman 81**: existence of the uniform value v_∞
- **Bravo and O.-B 2017**: solve the unknown stochastic game

Non zero-sum

- **Takahashi 62, Fink 64:** existence of stationary equilibria in Γ_λ

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series
- **Mertens and Neyman 81**: existence of the uniform value v_∞
- **Bravo and O.-B 2017**: solve the unknown stochastic game

Non zero-sum

- **Takahashi 62, Fink 64:** existence of stationary equilibria in Γ_λ
- **Sorin 86**: E_λ does not converge to E_∞ (the Paris Match)

Selected results on stochastic games

Zero-sum

- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series
- **Mertens and Neyman 81**: existence of the uniform value v_∞
- **Bravo and O.-B 2017**: solve the unknown stochastic game

Non zero-sum

- **Takahashi 62, Fink 64:** existence of stationary equilibria in Γ_λ
- **Sorin 86**: E_λ does not converge to E_∞ (the Paris Match)
- **Mertens-Parthasarathy 87**: stationary equilibria in Γ_λ (non finite)

Selected results on stochastic games

Zero-sum

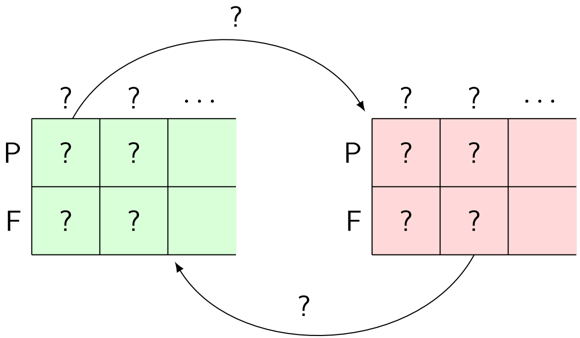
- **Shapley 53**: solves Γ_λ (fixed discount factor)
- **Blackwell and Ferguson 68**: solve the undiscounted Big Match
- **Bewley and Kohlberg 76**: convergence of the values v_λ as $\lambda \to 0$ and optimal strategies as Puiseux series
- **Mertens and Neyman 81**: existence of the uniform value v_∞
- **Bravo and O.-B 2017**: solve the unknown stochastic game

Non zero-sum

- **Takahashi 62, Fink 64:** existence of stationary equilibria in Γ_λ
- **Sorin 86**: E_λ does not converge to E_∞ (the Paris Match)
- **Mertens-Parthasarathy 87**: stationary equilibria in Γ_λ (non finite)
- **Vieille 00**: E_∞ is non empty for $N = 2$ (open for $N > 2$)

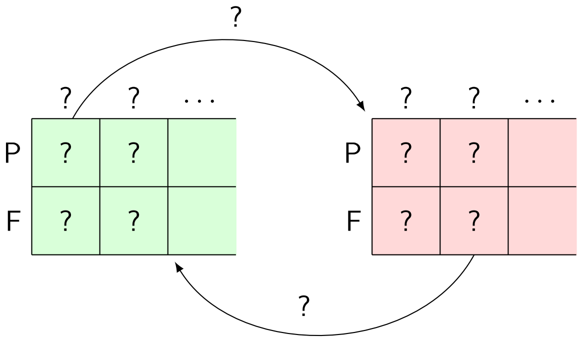
Unknown stochastic game (Bravo and O.-B.)

Stochastic game + unknown game



Unknown stochastic game (Bravo and O.-B.)

Stochastic game + unknown game



During the game player i observes $g^i(a_1)$, $g^i(a_2)$, $g^i(a_3)\ldots$

The player does not know anything else...

Our main result

Theorem (Bravo and O.-B. 17)

Let $\Gamma = (N, S, A, g, q)$ be a stochastic game and let v^i be the security level of player i. If q is ergodic, then Player i can guarantee v^i with minimal information, i.e. there exists σ^i such that for all strategy $\sigma^{-i} \in \Sigma^{-i}$

$$\liminf_{n \to +\infty} \mathbb{E}_{\sigma,q}\left[\frac{1}{n}\sum_{m=1}^{n} g^i(s_m, a_m)\right] \geq v^i$$

Our main result

Theorem (Bravo and O.-B. 17)

Let $\Gamma = (N, S, A, g, q)$ be a stochastic game and let v^i be the security level of player i. If q is ergodic, then Player i can guarantee v^i with minimal information, i.e. there exists σ^i such that for all strategy $\sigma^{-i} \in \Sigma^{-i}$

$$\liminf_{n \to +\infty} \mathbb{E}_{\sigma,q} \left[\frac{1}{n} \sum_{m=1}^{n} g^i(s_m, a_m) \right] \geq v^i$$

Corollary (Bravo and O.-B. 17)

If $\Gamma = (S, A, B, g, q)$ is a zero-sum stochastic game with value v, such that q is ergodic, then player 1 can guarantee v with minimal information.

Idea of the proof (1)

What can a player do ?

- First, he observes an initial state and his action set: he can play as in the unknown game (Auer et al. 1995)

Idea of the proof (1)

What can a player do ?

- First, he observes an initial state and his action set: he can play as in the unknown game (Auer et al. 1995)
- At some point, he reaches a new state and learns his new action set: he can play as in an unknown stochastic game (how !?) with 2 states

Idea of the proof (1)

What can a player do ?

- First, he observes an initial state and his action set: he can play as in the unknown game (Auer et al. 1995)
- At some point, he reaches a new state and learns his new action set: he can play as in an unknown stochastic game (how !?) with 2 states
- An so on... Under ergodicity assumptions, he will eventually reach all the states.

Idea of the proof (1)

What can a player do ?

- First, he observes an initial state and his action set: he can play as in the unknown game (Auer et al. 1995)
- At some point, he reaches a new state and learns his new action set: he can play as in an unknown stochastic game (how !?) with 2 states
- An so on... Under ergodicity assumptions, he will eventually reach all the states.

Solution: construct a strategy based on an auxiliary unknown game with infinite action sets and noise (X, Y, g) (the player plays x and observes $g(x, y) + U$, for some noise U) that has value v^i

Idea of the proof (2)

The two crucial points are:

- Extend Auer et al. 1995 to a more general unknown game (X, Y, g) where X is compact and $x \mapsto g(x, y)$ uniformly continuous on y
- Reduce the stochastic game to a fixed unknown game (X, Y, g) with X compact, $x \mapsto g(x, y)$ uniformly continuous and satisfying $\mathrm{val}(X, Y, g) = v^i$.

Idea of the proof (3)

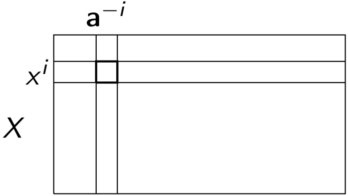
The auxiliary game (X, Y, h) is as follows:

- $X : \Delta(A)^S$ is the set of stationary strategies and $Y = (A^{-i})^{\mathbb{N}}$

Idea of the proof (3)

The auxiliary game (X, Y, h) is as follows:

- $X : \Delta(A)^S$ is the set of stationary strategies and $Y = (A^{-i})^{\mathbb{N}}$



- For each T the payoff is h_T given by (where $a_t^i \sim x^i$)

$$h_T^i(\sigma, \mathbf{a}^{-i}) = \frac{1}{T} \sum_{t=1}^{T} g^i(s_t, a_t)$$

- The value of (X, Y, h_T) approaches v^i as $T \to +\infty$

Perspectives

- The ergodicity assumption can be refined
- Study other, more general dynamic games, under the minimal information setup (games which we can solve at noon...)
- Study the speed convergence rates and computability issues

Thank you for your attention