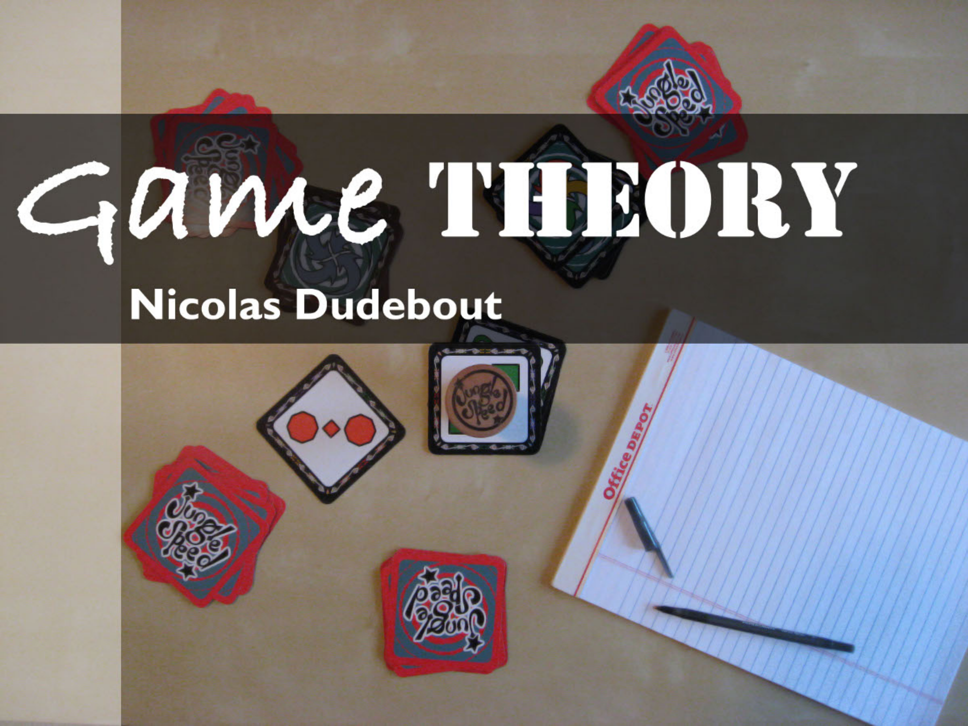
Game THEORY

Nicolas Dudebout



Previously in *Game Theory*

Previously in *Game Theory*

- decision makers:
  - choices
  - preferences

Previously in *Game Theory*

- decision makers:
  - choices
  - preferences
- solution concepts:
  - best response
  - Nash equilibrium

Rock, paper, scissors

Rock, paper, scissors

	R	P	S
R	$0, 0$	$-1, 1$	$1, -1$
P	$1, -1$	$0, 0$	$-1, 1$
S	$-1, 1$	$1, -1$	$0, 0$

Learning in games

Learning in games
Repeated games

Learning in games

Best Response learning

Best Response learning

1. Guess what the opponent(s) will play

Best Response learning

1. Guess what the opponent(s) will play
2. Play a Best Response to that guess

Best Response learning

1. Guess what the opponent(s) will play
2. Play a Best Response to that guess
3. Observe the play

Best Response learning

1. Guess what the opponent(s) will play
2. Play a Best Response to that guess
3. Observe the play
4. Update the guess

BR learning: Cournot dynamics

BR learning: Cournot dynamics

Guess = last action played

BR learning: Cournot dynamics

Guess = last action played

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

BR learning: Cournot dynamics

Guess = last action played

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

	R	P	S
R	$0, 0$	$-1, 1$	$1, -1$
P	$1, -1$	$0, 0$	$-1, 1$
S	$-1, 1$	$1, -1$	$0, 0$

BR learning: Fictitious play

BR learning: Fictitious play

Guess $=$ empirical distribution of play

BR learning: Fictitious play

Guess = empirical distribution of play

	R	P	S
R	$0, 0$	$-1, 1$	$1, -1$
P	$1, -1$	$0, 0$	$-1, 1$
S	$-1, 1$	$1, -1$	$0, 0$

BR learning: Fictitious play

Guess = empirical distribution of play

	R	P	S
R	$0, 0$	$-1, 1$	$1, -1$
P	$1, -1$	$0, 0$	$-1, 1$
S	$-1, 1$	$1, -1$	$0, 0$

	L	C	R
U	$0, 0$	$0, 1$	$1, 0$
M	$1, 0$	$0, 0$	$0, 1$
D	$0, 1$	$1, 0$	$0, 0$

Evolutionary learning

Evolutionary learning

Action set: A
Utility function: u

Evolutionary learning

Action set: A
Utility function: u

$p \in \Delta(A), k \in A$
$\dot{p}_k = p_k \left(u(k,p) - u(p,p)\right)$

Battle of the Sexes

Battle of the Sexes

	O	F
O	$3, 2$	$0, 0$
F	$0, 0$	$2, 3$

Correlated equilibrium (CE)

Correlated equilibrium (CE)

$a^* \in A = \prod_i A_i$ is a NE:

$$\forall i, \forall a'_i, u_i(a^*_i, a^*_{-i}) \geq u_i(a'_i, a^*_{-i})$$

Correlated equilibrium (CE)

$a^* \in A = \prod_i A_i$ is a NE:

$$\forall i, \forall a_i', u_i(a_i^*, a_{-i}^*) \geq u_i(a_i', a_{-i}^*)$$

$\alpha \in \prod_i \Delta(A_i)$ is a NE: $\forall i, \forall a_i, \forall a_i',$

$$\sum_{a_{-i}} u_i(a_i, a_{-i})\alpha(a) \geq \sum_{a_{-i}} u_i(a_i', a_{-i})\alpha(a)$$

Correlated equilibrium (CE)

$a^* \in A = \prod_i A_i$ is a NE:

$$\forall i, \forall a_i', u_i(a_i^*, a_{-i}^*) \geq u_i(a_i', a_{-i}^*)$$

$\alpha \in \prod_i \Delta(A_i)$ is a NE: $\forall i, \forall a_i, \forall a_i',$

$$\sum_{a_{-i}} u_i(a_i, a_{-i})\alpha(a) \geq \sum_{a_{-i}} u_i(a_i', a_{-i})\alpha(a)$$

$\pi \in \Delta(A)$ is a CE: $\forall i, \forall a_i, \forall a_i',$

$$\sum_{a_{-i}} u_i(a_i, a_{-i})\pi(a) \geq \sum_{a_{-i}} u_i(a_i', a_{-i})\pi(a)$$

No regret learning

No regret learning

$$u_i(k, a_{-i}) - u_i(j, a_{-i})$$

No regret learning

$$u_i(k, a_{-i}) - u_i(j, a_{-i})$$

$$R^i_{jk}(t) = \sum_{\tau=0:a_i(\tau)=j}^{t} u_i(k, a_{-i}(\tau)) - u_i(j, a_{-i}(\tau))$$

No regret learning

$$u_i(k, a_{-i}) - u_i(j, a_{-i})$$

$$R^i_{jk}(t) = \sum_{\tau=0:a_i(\tau)=j}^{t} u_i(k, a_{-i}(\tau)) - u_i(j, a_{-i}(\tau))$$

Regret matching converges to the correlated equilibria set.

Learning in games

Learning in games

- Best response

Learning in games

- Best response
- Replicator dynamics

Learning in games

- Best response
- Replicator dynamics
- No regret

Repeated games

Markov Decision Process (MDP)

Markov Decision Process (MDP)

state space X

action space U

transition $P : X \times U \to \Delta(X)$

reward $r : X \times U \to \mathbb{R}$

discount factor $\delta \in [0, 1]$

Markov Decision Process (MDP)

state space X

action space U

transition $P : X \times U \to \Delta(X)$

reward $r : X \times U \to \mathbb{R}$

discount factor $\delta \in [0, 1]$

$$U(x(\cdot), u(\cdot)) = \sum_{t=0}^{+\infty} \delta^t r(x(t), u(t))$$

MDP (continued)

history $\mathcal{H} \in \prod(X, U)$
policy $\pi : \mathcal{H} \to \Delta(U)$

MDP (continued)

history $\mathcal{H} \in \prod(X, U)$
policy $\pi : \mathcal{H} \to \Delta(U)$
$V^\pi(x_0) = \mathbb{E}_\pi \left[U(x(\cdot), u(\cdot)) \right]$

MDP (continued)

history $\mathcal{H} \in \prod(X, U)$
policy $\pi : \mathcal{H} \to \Delta(U)$
$V^\pi(x_0) = \mathbb{E}_\pi \left[U(x(\cdot), u(\cdot)) \right]$

$$V(x_0) = \max_\pi V^\pi(x_0)$$

Principle of Optimality

Bellman's equation:

$$V(x_0) = \max_{u_0} \left[r(x_0, u_0) + \delta V(P(x_0, u_0)) \right]$$

Dynamic Programming

Solving the MDP:

Dynamic Programming

Solving the MDP:

- knowing P: value iteration

Dynamic Programming

Solving the MDP:

- knowing P: value iteration
- not knowing P: online learning

Repeated game

Repeated game

Game $(\mathcal{I}, \prod_i A_i, \prod_i u_i)$

Repeated game

Game $(\mathcal{I}, \prod_i A_i, \prod_i u_i)$
Discount factor δ

$$U_i(a(\cdot)) = \sum_{t=0}^{+\infty} \delta^t u_i(a(t))$$

Repeated game

Game $(\mathcal{I}, \prod_i A_i, \prod_i u_i)$
Discount factor δ

$$U_i(a(\cdot)) = \sum_{t=0}^{+\infty} \delta^t u_i(a(t))$$

Strategy $\sigma : \mathcal{H} \to \prod_i \Delta(A_i x)$

$$V_i(\sigma) = \mathbb{E}_\sigma \left[U_i(a(\cdot))\right]$$

Nash equilibrium

Player i:

- choices σ_i
- utility V_i

Nash equilibrium

Player i:

- choices σ_i
- utility V_i

Nash equilibrium is not strong enough!
(Explanation on the whiteboard $\Rightarrow$)

Information structure

Information structure

- perfect
- imperfect

Information structure

- perfect
- imperfect

- public
- private (beliefs)

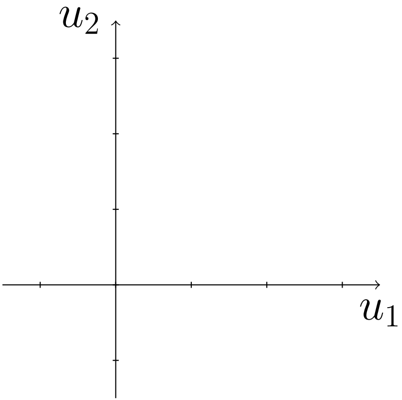
Folk theorem

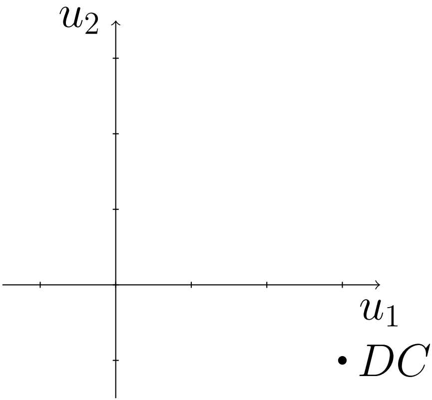
Any feasible, strictly individually rational payoff can be sustained by a sequentially rational equilibrium.

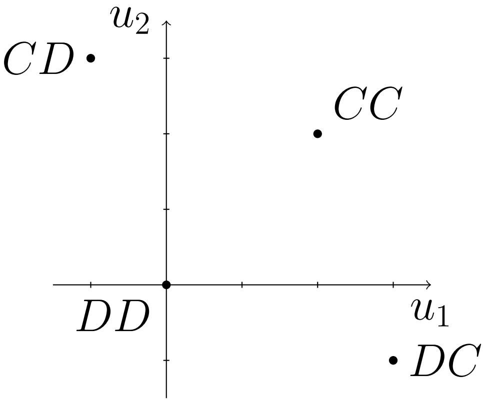
Folk theorem

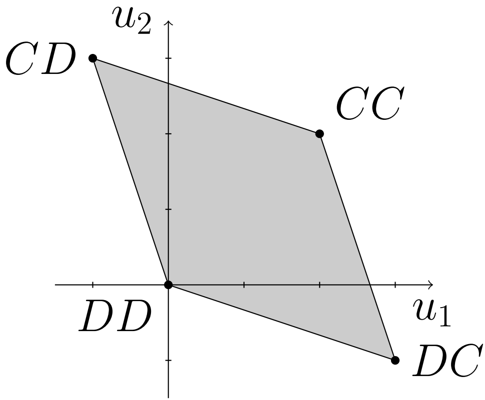
Any feasible, strictly individually rational payoff can be sustained by a sequentially rational equilibrium.

Holy grail for repeated games.

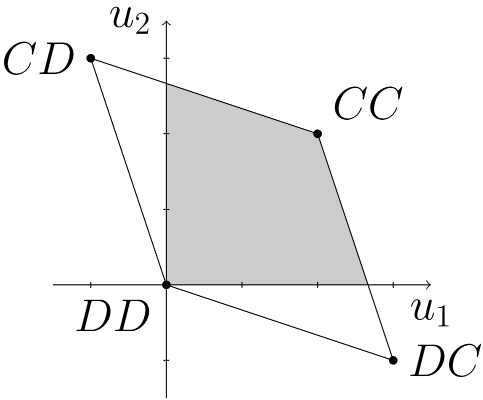

u_2
u_1


u_2
u_1
DC


u_2
CD
CC
u_1
DD
DC



u_2
CD
CC
DD
u_1
DC


u_2
CD
CC
DD
u_1
DC

Research

Weakly belief-free equilibria

Characterization of repeated games with correlated equilibria.

Repeated games

Repeated games

- Dynamic programming

Repeated games

- Dynamic programming
- Repeated games

Repeated games

- Dynamic programming
- Repeated games
- Folk theorem

Learning in games

Learning in games
Repeated games

Questions, Comments