INFO-F-409
Learning dynamics

Learning, evolutionary game theory and the evolution of co-operation

T. Lenaerts and A. Nowé
MLG, Université Libre de Bruxelles and
AI-lab, Vrije Universiteit Brussel

Summary

- What? Why?
- Rational choice
- Strategic games
- Nash Equilibrium
- Best
- Dominance
- Mixed strategies
- Mixed-strategy Nash Equilibria
- Support finding
- Lemke-Howson algorithm
- Extensive-form games
- sub-game perfect equilibrium
- Simultaneous moves
- Chance moves
- Bayesian games
- Assignment 1

The formation of agents' beliefs

Now that we can determine the Nash and sub-game perfect equilibria ...

How can we reach them?

Which equilibrium is preferred ?

The formation of agents' beliefs

Can we expect that the equilibrium will be reached ?

Players could chose their action from an introspective analysis of the game : removing dominated strategies

Learning the beliefs about the other player in response of the information she receives :

1. Best response dynamics
2. Fictitious play
3. Stimulus-response or reinforcement learning
4. Evolutionary or cultural dynamics

Levels of learning

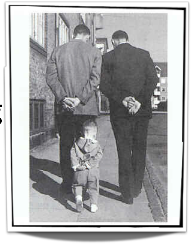
innate

reflex actions

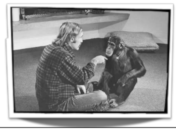
imprinting (specific and irreversible)

Conditioning

Observational and imitative learning



Teaching



learned

Conditioning

Scene from the Big Bang Theory (S03E03, The Gothowitz Deviation)

Conditioning

Scene from the Big Bang Theory (S03E03, The Gothowitz Deviation)

Best-response dynamics

In the first period, choose a best response to an arbitrary deterministic belief about the other players' actions

In every period after the first, choose the best response to the other players' actions in the previous round

An action profile that remains the same over time is a pure Nash equilibrium of the game

Best-response dynamics

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

Depending on the prior beliefs these dynamics may not converge

Take for instance the Battle of the sexes, which has 3 equilibria ((1,0),(1,0)), ((0,1),(0,1)) and ((2/3,1/3),(1/3, 2/3))

8-1

Best-response dynamics

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

Depending on the prior beliefs these dynamics may not converge

Take for instance the Battle of the sexes, which has 3 equilibria ((1,0),(1,0)), ((0,1),(0,1)) and ((2/3,1/3),(1/3, 2/3))

BELIEF	A plays	B plays
prior	B	B
1	B	B
2	B	B
...	...	...

8-2

Best-response dynamics

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

Depending on the prior beliefs these dynamics may not converge

Take for instance the Battle of the sexes, which has 3 equilibria ((1,0),(1,0)), ((0,1),(0,1)) and ((2/3,1/3),(1/3, 2/3))

BELIEF	A plays	B plays
prior	B	B
1	B	B
2	B	B
...	...	...

BELIEF	A plays	B plays
prior	S	S
1	S	S
2	S	S
...	...	...

8-3

Best-response dynamics

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

Depending on the prior beliefs these dynamics may not converge

Take for instance the Battle of the sexes, which has 3 equilibria ((1,0),(1,0)), ((0,1),(0,1)) and ((2/3,1/3),(1/3, 2/3))

BELIEF	A plays	B plays
prior	B	B
1	B	B
2	B	B
...	...	...

BELIEF	A plays	B plays
prior	S	S
1	S	S
2	S	S
...	...	...

BELIEF	A plays	B plays
prior	S	B
1	B	S
2	S	B
...	...	...

8-4

Fictitious play

Every agent starts with an arbitrary probabilistic belief about the other players actions.

In the first round she chooses a BR to this prior probabilistic belief and observes the other player's actions, say A.

she changes here belief so that A gets probability 1

In the second round, she produces a best response to this belief and observes the other player's action, say B

she changes here belief to one that assigns 1/2 to action A and 1/2 to action B

In the third round ...

Fictitious play

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

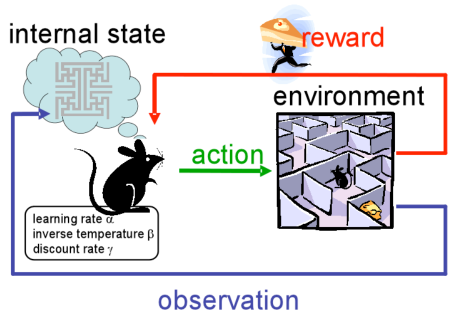
	BELIEF				
	A plays	B>A	B plays	A>B	
prior		(1,0)		(0,1)	
1	S	(1,1)	B	(1,1)	TOTAL = 2
2	S	(1,2)	S	(1,2)	TOTAL = 3
3	S	(1,3)	S	(1,3)	TOTAL = 4
4	S	(2,3)	B	(1,4)	TOTAL = 5
5	S	(2,4)	S	(1,5)	TOTAL = 6
6	S	(2,5)	S	(1,6)	TOTAL = 7
7	...	...	...	...	

Fictitious play

So in any period, the agent adopts the belief that her opponent is using a mixed strategy in which the probability of each action is proportional to the frequency with which her opponent has chosen that action in the previous rounds

The process converges to a mixed strategy Nash equilibrium from initial beliefs

Stimulus-response learning



Stimulus-response learning

Stochastic dynamic models of individual behavior ...

Bush, R. R., & Mosteller, F. (1951). **A mathematical model for simple learning.** Psychological review, 58(5), 313–323.

Roth, A. E., & Erev, I. (1995). **Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term.** Games and Economic Behavior, 8(1), 164–212.

Erev, I., & Roth, A. E. (1998). **Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria.** The American Economic Review, 88(4), 848–881.

Stimulus-response learning

Take for instance the model proposed by Roth and Erev (1995)

A player is defined by:

A **propensity score** $q_{Ak}(t)$, which expresses the propensity of player A to play action k at time t

A **probability function** $p_{Ak}(t) = q_{Ak}(t)/\sum_j q_{Aj}(t)$, which expresses the probability of A to play action k at time t

An **update function** $q_{Ak}(t+1) = q_{Ak}(t) + x$, where x is the payoff from the interaction. The other actions $q_{Aj}(t)$ remain the same.

Hence actions with a higher probability are more likely to be played (**Law of effect**)

Aim was to design a model that fits psychological literature

Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*1*	*1/2*
Stravinsky	*1*	*1/2*

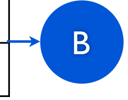
player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*1*	*1/2*
Stravinsky	*1*	*1/2*

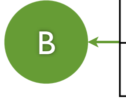
Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*1*	*1/2*
Stravinsky	*1*	*1/2*





player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*1*	*1/2*
Stravinsky	*1*	*1/2*

Stimulus-response learning

	Bach	Strav.
Bach	1 2	0 0
Strav.	0 0	2 1

Consider again the Battle of the sexes:

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*1*	*1/2*
Stravinsky	*1*	*1/2*

B

2

B

1

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*1*	*1/2*
Stravinsky	*1*	*1/2*

Stimulus-response learning

	Bach	Strav.
Bach	1 2	0 0
Strav.	0 0	2 1

Consider again the Battle of the sexes:

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

	Bach	Strav.
Bach	1 2	0 0
Strav.	0 0	2 1

Consider again the Battle of the sexes:

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

	Bach	Strav.
Bach	1 2	0 0
Strav.	0 0	2 1

Consider again the Battle of the sexes:

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

B

S

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

B

0

S

0

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

S

S

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	*3*	*3/4*
Stravinsky	*1*	*1/4*

S

1

S

2

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	*2*	*2/3*
Stravinsky	*1*	*1/3*

Stimulus-response learning

Consider again the Battle of the sexes:

	Bach	Strav.
Bach	2, 1	0, 0
Strav.	0, 0	1, 2

player 1	$q_{1k}(t)$	$p_{1k}(t)$
Bach	3	3/5
Stravinsky	2	2/5

player 2	$q_{2k}(t)$	$p_{2k}(t)$
Bach	2	2/5
Stravinsky	3	3/5

Continue until convergence

Stimulus-response learning

Three extensions were introduced into this model:

A **cutoff parameter** μ which ensure that $q_{Ak}(t)$ and $p_{Ak}(t)$ can become zero in finite time : when $p_{Ak}(t) \leq \mu$, $q_{Ak}(t)=p_{Ak}(t)=0$

An **error/exploration parameter** ε which prevents a probability $p_{Ak}(t)$ can become zero if it is close to a successful strategy: $q_{Ak}(t+1)=q_{Ak}(t)+(1\text{-}\varepsilon)x$ for the successful strategy and $q_{Aj}(t+1)=q_{Aj}(t)+\varepsilon x$ for *the adjacent strategies*

An **forgetting parameter** φ which gradually reduces the importance of each propensity $q_{Ak}(t)$ over time by multiplying each propensity by $(1\text{-}\varphi)$.

More details on reinforcement learning by Prof. Nowé