

# Evolutionary game dynamics and convergence to equilibria: a survey

Yannick Viossat (Ceremade, Université Paris-Dauphine)

AlgoGT 2011, Grenoble, 20-21 June 2011

# What is evolutionary game theory ?

Traditional game theory :

- rational agents,
- fully understand the game, unbounded computation abilities
- all this common knowledge.

Evolutionary game theory

- boundedly rational agents
- need not fully understand the game
- use rules of thumb (imitation, trial and error, natural selection,...)

# Two kinds of models

## Population dynamics

- large population of agents
- from time to time, randomly drawn agents interact
- state variable : current population mean behavior (no memory)

## Learning

- two or a few agents
- play repeatedly against each other (need not understand it)
- state variable : simple statistic of past history ;  
e.g. empirical frequencies of opponent's actions.

- 1 Are irrational behaviors (dominated strategies) eliminated?
- 2 Does the evolutionary process lead to some equilibrium?
- 3 If so, which one?

Here, topic 2.

- 1 Population dynamics :
  - framework, examples
  - convergence results
  - nonconvergence results
- 2 Learning :
  - no-regret dynamics and convergence to correlated equilibria
  - learning to be often in equilibrium à la Young

# Population dynamics : framework and notation

- single large population
- randomly drawn agents play a two-player symmetric game
- $n$  possible pure strategies :  $1, 2, \dots, n$
- Agents play pure strategies, but may switch
- $x_i(t)$  : frequency of strategy  $i$  at time  $t$
  
- $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$  : state variable
- state space :  $\{x = (x_1, \dots, x_n) \in \mathbb{R}_+^n, \sum_i x_i = 1\}$

Population dynamics :  $\dot{x} = f(x, \text{payoffs})$  (no memory).

Payoffs :

- $u(i, j)$  : payoff of strategy  $i$  against strategy  $j$
- expected payoff of strategy  $i$  against  $x$  :  $u(i, x) = \sum_j x_j u(i, j)$
- mean payoff :  $x \cdot u(x)$ , where  $u(x) = (u(1, x), \dots, u(n, x))$
- $x$  symmetric Nash equilibrium if  $x \cdot u(x) \geq y \cdot u(x) \forall y \in S_n$

# Example of dynamics

- Best-reply dynamics (Gilboa & Matsui, 91)

$$\dot{x}(t) = b(t) - x(t) \quad \text{with } b(t) \text{ best-response to } x(t)$$

↪ Evolution towards best-replies to current behavior

- Replicator Dynamics (Taylor & Jonker, 78) :

$$\dot{x}_i = x_i [u(i, x) - x \cdot u(x)]$$

↪ Growth-rate of a strategy = own payoff – average payoff



# Classes of dynamics

- monotone dynamics : a higher payoff yields a higher growth-rate
- adaptive dynamics : mean-payoff increases against a stationary environment
- etc.

Link between outcomes of such dynamics and Nash equilibria ?

# Convergence results (roughly stated)

**Evolutionary Folk Theorem** : if a “reasonable” dynamics converges to a point, then this point is a Nash equilibrium

## Convergence in special classes of games

- Common interest games, potential games : all adaptive dynamics
- Dominance solvable games : best-reply and convex monotone dynamics
- Games with self-defeating externalities : best-reply, replicator, others
- Games with strategic complementarities : best-reply (and many others when equilibrium unique)

For precise statements, see Sandholm (2011).

# Divergence results

- in versions of Rock-Paper-Scissors, many dynamics cycle away from the unique equilibrium. E.g. replicator, best-reply.
- cycling behavior universal : occurs for any smooth adaptive dynamics
- even cycling between strategies with zero probability in all equilibria
- for monotone dynamics, similar results for correlated equilibria (an extension of Nash equilibria)
- for replicator and best-reply, in some games, for almost all initial conditions, only strategies with proba zero in all equilibria survive

Omitted (see next talk)

- no-regret dynamics and convergence to correlated equilibria
- learning to be often in equilibrium

# No-regret dynamics : framework

A player plays repeatedly a finite 2-player game against some opponent

- at each step  $t$ , he chooses action  $i_t$  and receives payoff  $u(i_t, j_t)$

- average payoff until step  $T$  :  $\frac{1}{T} \sum_1^T u(i_t, j_t)$

- payoff had he always played  $i$  :  $\frac{1}{T} \sum_1^T u(i, j_t)$

Regret for not having played  $i$  = difference between these quantities :

$$R_T(i) = \frac{1}{T} \sum_1^T u(i, j_t) - \frac{1}{T} \sum_1^T u(i_t, j_t)$$

The player has asymptotically no-regret if  $\limsup_{T \rightarrow +\infty} R_T(i) \leq 0, \forall i$

# No-regret procedure

No-regret procedure : leads to no-regret for any opponent's strategy

Uncoupled procedure : does not depend on the opponent's payoffs

**Theorem 1** (Hannan, 1957) : there are uncoupled no-regret procedures.

E.g. **regret-matching** :

- if some regrets positive : play only strategies with positive regret, with probability proportional to the regret
- if all regrets nonpositive : play whatever

# Conditional regret (also called internal regret)

*Conditional regret* for not having played  $i$  instead of  $k$  :

increase in average payoff if whenever I played  $k$ , I had played  $i$  instead

No conditional regret : all the conditional regrets are nonpositive.

**Theorem 2** : there are uncoupled "no conditional regret" strategies



# Correlated equilibrium

An extension of Nash equilibrium to situations where players may condition their actions on payoff-irrelevant signals

Set of correlated equilibria : convex polytope containing the Nash eq.

No conditional regret for all players = correlated equilibrium

**Theorem 2'** : there are uncoupled procedures s.t. if all players use such procedures, the average play converges to the set of correlated equilibria

**Theorem 3** : there are no uncoupled procedures leading to Nash equilibria

# Comparison with population dynamics

- There are population dynamics leading to no-regret (e.g. replicator)
- but none are known to lead to no-conditional regret
- For monotone dynamics and others, this cannot be

Tentative explanation : not the same state space ; not the same memory

# Learning to be often in efficient equilibrium (Young)

- Agents characterized by a mood, a benchmark action and a benchmark payoff
- Four possible moods : content, discontent, hopeful, watchful
- Agents change their search behavior according to their mood
  - when content they experiment occasionally
  - when discontent they experiment frequently
- Hopeful and watchful are transitional moods

# Example of transitions - I

- Content-content :
  - when content an agent experiment with some small probability
  - if leads to higher payoff, adopt the experimental action and the experienced payoff as new benchmarks, and stays content
- Discontent-content :
  - when discontent, experiment
  - with fixed probability, adopt experimental action and experienced payoff as benchmarks, and become content

## Example of transitions - II

Transitions involving hopeful and watchful moods : triggered by changes in payoffs without experimenting (i.e. due to others).

E.g. if payoff increases without experimenting, a content agent becomes hopeful, and then :

- becomes content again if payoff stays above benchmark
- becomes watchful otherwise

Let  $\epsilon > 0$ . There are uncoupled learning rules of this kind such that, in any game that satisfies mild conditions and has a pure Nash equilibrium, experienced play is an efficient equilibrium more than  $1 - \epsilon$  of the time.

Young recently suggested that modifications of these rules should allow to lead to a welfare maximizing state, and not only to an equilibrium.

Appealing for applications where we can choose the agents learning rules

## **Population dynamics :**

Hofbauer, J. and K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge University Press, 1998

Sandholm W. Population Games and Evolutionary Dynamics, MIT Press, 2011

Weibull, J., Evolutionary Game Theory, MIT Press, 1995

## **No regret dynamics :**

Babichenko, Y., "Completely uncoupled dynamics and Nash equilibria," working paper, 2010

Hart S., "Adaptive heuristics", Econometrica 73, 1401-1430, 2005

## **Learning to be often in equilibrium :**

Pradelski, B. and H. P. Young, "Learning efficient Nash equilibria in distributed systems", working paper, 2010

Young, H. P., Strategic Learning and Its Limits, Oxford University Press, 2004