SIKS tutorial "Agent Systems"

Multi-agent learning

Gerard Vreeswijk

Intelligent Systems Group, Computer Science Department
Faculty of Sciences, Utrecht University, The Netherlands

December 9, 2013

Objective of this presentation

Objective

Objective of this presentation

Objective

To consider processes of **adaptation**

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games**

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games** with **real-valued actions**.

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games** with **real-valued actions**.

- Choose how much lemonade to produce.

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games** with **real-valued actions**.

- Choose how much lemonade to produce.
- Or bandwidth (think competition between broadband providers).

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games** with **real-valued actions**.

- Choose how much lemonade to produce.
- Or bandwidth (think competition between broadband providers).
- Or personal attention (think competition between social bots).

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games** with **real-valued actions**.

- Choose how much lemonade to produce.
- Or bandwidth (think competition between broadband providers).
- Or personal attention (think competition between social bots).
- Players **adapt** to each other's actions.

Objective of this presentation

Objective

To consider processes of **adaptation** in **repeated two-player games** with **real-valued actions**.

- Choose how much lemonade to produce.
- Or bandwidth (think competition between broadband providers).
- Or personal attention (think competition between social bots).
- Players **adapt** to each other's actions.
- For some modes of adaptation, interesting dynamics will occur.

Outline

1. Introduction
   - Multi-agent learning (MAL)
   - Teaching
2. Cournot dynamics
   - Cournot competition
   - Cournot equilibrium
3. Alternative Cournot dynamics
   - Work of David Rand (1978)
   - Work of Tönu Puu (1991)
   - Conclusions

Outline

Outline

What is multi-agent learning?

Traditional machine learning:

What is multi-agent learning?

Traditional machine learning:

- Learning **offline**: statistics / data mining.

What is multi-agent learning?

Traditional machine learning:

- Learning **off**line: statistics / data mining.
- Learning **on**line in a non-influenceable environment.

What is multi-agent learning?

Traditional machine learning:

- Learning **off**line: statistics / data mining.
- Learning **on**line in a non-influenceable environment.
- Learning **on**line in a changing but non-influenceable environment.

What is multi-agent learning?

Traditional machine learning:

- Learning **off**line: statistics / data mining.
- Learning **on**line in a non-influenceable environment.
- Learning **on**line in a changing but non-influenceable environment.

Multi-agent learning:

What is multi-agent learning?

Traditional machine learning:

- Learning **off**line: statistics / data mining.
- Learning **on**line in a non-influenceable environment.
- Learning **on**line in a changing but non-influenceable environment.

Multi-agent learning:

- Vs. one other learning agent (playing $1 - 1$).

What is multi-agent learning?

Traditional machine learning:

- Learning **off**line: statistics / data mining.
- Learning **on**line in a non-influenceable environment.
- Learning **on**line in a changing but non-influenceable environment.

Multi-agent learning:

- Vs. one other learning agent (playing $1 - 1$).
- Vs. multiple learning agents (playing $n - n$).

What is multi-agent learning?

Traditional machine learning:

- Learning **off**line: statistics / data mining.
- Learning **on**line in a non-influenceable environment.
- Learning **on**line in a changing but non-influenceable environment.

Multi-agent learning:

- Vs. one other learning agent (playing $1 - 1$).
- Vs. multiple learning agents (playing $n - n$).
- Vs. very many other learning agents. (Large populations, but playing $1 - 1$.)

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Adaptation =

learning (following)

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Adaptation =
learning (following) as well as **teaching** (exercising influence).

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Adaptation =
learning (following) as well as **teaching** (exercising influence).

- Study the adaptation to other's behaviour (**descriptive**).

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Adaptation =
learning (following) as well as **teaching** (exercising influence).

- Study the adaptation to other's behaviour (**descriptive**).
- Prescribe the adaptation to other's behaviour (**normative**).

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Adaptation =
learning (following) as well as **teaching** (exercising influence).

- Study the adaptation to other's behaviour (**descriptive**).
- Prescribe the adaptation to other's behaviour (**normative**).
- Formalise and study **emergence** in multi-agent systems.

Why multi-agent learning?

Goal of multi-agent learning

To understand and / or prescribe processes of (mostly) social adaptation in systems of artificial agents.

Adaptation =
learning (following) as well as **teaching** (exercising influence).

- Study the adaptation to other's behaviour (**descriptive**).
- Prescribe the adaptation to other's behaviour (**normative**).
- Formalise and study **emergence** in multi-agent systems.
- Explain how Nash equilibria may come about.

Outline

Teaching

A game in normal form

	L	R
T	$1, 0$	$3, 2$
B	$2, 1$	$4, 0$

Teaching

A game in normal form

	L	R
T	$1, 0$	$3, 2$
B	$2, 1$	$4, 0$

- Row player has a dominating strategy: B.

Teaching

A game in normal form

	L	R
T	$1, 0$	$3, 2$
B	$2, 1$	$4, 0$

- Row player has a dominating strategy: B.
- Pure Nash equilibrium: (B, L) with payoff profile $(2, 1)$.

Teaching

A game in normal form

	L	R
T	$1,0$	$3,2$
B	$2,1$	$4,0$

- Row player has a dominating strategy: B.
- Pure Nash equilibrium: (B, L) with payoff profile $(2, 1)$.
- However: action profile (T, R) with payoff profile $(3, 2)$ Pareto dominates the equilibrium.

Teaching

A game in normal form

	L	R
T	$1, 0$	$3, 2$
B	$2, 1$	$4, 0$

- Row player has a dominating strategy: B.
- Pure Nash equilibrium: (B, L) with payoff profile $(2, 1)$.
- However: action profile (T, R) with payoff profile $(3, 2)$ Pareto dominates the equilibrium.
- Both can achieve the Pareto optimum if the row player teaches T, and the column player recognises this, and follows.

Coordination game

Who should be teaching here?

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

Coordination game

Who should be teaching here?

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

- Intuition: left- vs. right-hand traffic.

Coordination game

Who should be teaching here?

	L	R
L	$1,1$	$0,0$
R	$0,0$	$1,1$

- Intuition: left- vs. right-hand traffic.
- Optimal profiles are (L, L) (UK) and (R, R) (the rest).

Coordination game

Who should be teaching here?

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

- Intuition: left- vs. right-hand traffic.
- Optimal profiles are (L, L) (UK) and (R, R) (the rest).
- Suppose start profile is (R, L).

Coordination game

Who should be teaching here?

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

- Intuition: left- vs. right-hand traffic.
- Optimal profiles are (L, L) (UK) and (R, R) (the rest).
- Suppose start profile is (R, L).
- If both teach:

$$(R, L) \to (R, L) \to (R, L) \to (R, L) \to \dots$$

Coordination game

Who should be teaching here?

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

- Intuition: left- vs. right-hand traffic.
- Optimal profiles are (L, L) (UK) and (R, R) (the rest).
- Suppose start profile is (R, L).
- If both teach:
$$(R, L) \to (R, L) \to (R, L) \to (R, L) \to \dots$$
- If both follow:
$$(R, L) \to (L, R) \to (R, L) \to (L, R) \to \dots$$

Forms of multi-agent learning

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.
- Regret-matching.

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.
- Regret-matching.
- Rational learning (a.k.a. Bayesian learning).

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.
- Regret-matching.
- Rational learning (a.k.a. Bayesian learning).
- Learning through evolution.

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.
- Regret-matching.
- Rational learning (a.k.a. Bayesian learning).
- Learning through evolution.
- Win-stay loose-shift / Win or Learn Fast (WoLF) priciples.

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.
- Regret-matching.
- Rational learning (a.k.a. Bayesian learning).
- Learning through evolution.
- Win-stay loose-shift / Win or Learn Fast (WoLF) priciples.
- Maintaining and rejecting hypothesis.

Forms of multi-agent learning

- Rudimentary strategies: tit-for-tat, trigger, cooperative, unforgiving (Robert Axelrod's tournaments in 1980)
- Reinforcement learning.
- Fictitious play.
- Regret-matching.
- Rational learning (a.k.a. Bayesian learning).
- Learning through evolution.
- Win-stay loose-shift / Win or Learn Fast (WoLF) priciples.
- Maintaining and rejecting hypothesis.
- . . .

Outline

Cournot dynamics

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.
- Inspired by spring water duopoly.

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.
- Inspired by spring water duopoly.
- All agents produce the same homogeneous product.

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.
- Inspired by spring water duopoly.
- All agents produce the same homogeneous product.
- Price is a commonly known decreasing function of total output.

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.
- Inspired by spring water duopoly.
- All agents produce the same homogeneous product.
- Price is a commonly known decreasing function of total output.
- Agents compete in the quantities they produce.

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.
- Inspired by spring water duopoly.
- All agents produce the same homogeneous product.
- Price is a commonly known decreasing function of total output.
- Agents compete in the quantities they produce.
- They choose quantities independently and simultaneously.

Antoine Augustin Cournot (1801-1877)

Cournot dynamics

- Economic model of duopoly.
- Inspired by spring water duopoly.
- All agents produce the same homogeneous product.
- Price is a commonly known decreasing function of total output.
- Agents compete in the quantities they produce.
- They choose quantities independently and simultaneously.
- Agents maximise profit given their competitors' decisions.

Antoine Augustin Cournot (1801-1877)

Outline

1. Introduction
   - Multi-agent learning (MAL)
   - Teaching
2. Cournot dynamics
   - Cournot competition
   - Cournot equilibrium
3. Alternative Cournot dynamics
   - Work of David Rand (1978)
   - Work of Tönu Puu (1991)
   - Conclusions

Cournot competition

Assumptions:

Cournot competition

Assumptions:

- **Agent 1** and **Agent 2**.

Cournot competition

Assumptions:

- **Agent 1** and **Agent 2**.
- Choose real-valued quantities x and y of the same good to produce (simultaneously).

Cournot competition

Assumptions:

- **Agent 1** and **Agent 2**.
- Choose real-valued quantities x and y of the same good to produce (simultaneously).
- These quantities are not revealed.

Cournot competition

Assumptions:

- **Agent 1** and **Agent 2**.
- Choose real-valued quantities x and y of the same good to produce (simultaneously).
- These quantities are not revealed.
- The **revenue** per unit is a monotone decreasing function of $x + y$.

Cournot competition

Assumptions:

- **Agent 1** and **Agent 2**.
- Choose real-valued quantities x and y of the same good to produce (simultaneously).
- These quantities are not revealed.
- The **revenue** per unit is a monotone decreasing function of $x + y$. Let us say:

$$r =_{Def} \max\{0, a - (x + y)\},$$

where $a > 0$ is some saturation level.

Cournot competition

Assumptions:

- **Agent 1** and **Agent 2**.
- Choose real-valued quantities x and y of the same good to produce (simultaneously).
- These quantities are not revealed.
- The **revenue** per unit is a monotone decreasing function of $x + y$. Let us say:

$$r =_{Def} \max\{0, a - (x + y)\},$$

  where $a > 0$ is some saturation level.
- There are **production costs** per unit. These are c, with $0 < c < a$.

Cournot competition

Profit:

Cournot competition

Profit:

- The **profit** for **Agent 1** per unit is therefore

$$\pi_1(x) = rx - cx$$

Cournot competition

Profit:

- The **profit** for **Agent 1** per unit is therefore

$$\begin{aligned}\pi_1(x) &= rx - cx \\ &= \begin{cases} (a-(x+y))x - cx & \text{if } x+y \leq a, \\ 0 - cx & \text{else.} \end{cases}\end{aligned}$$

Cournot competition

Profit:

- The **profit** for **Agent 1** per unit is therefore

$$
\begin{aligned}
\pi_1(x) &= rx - cx \\
&= \begin{cases} (a-(x+y))x - cx & \text{if } x+y \leq a, \\ 0 - cx & \text{else.} \end{cases} \\
&= \begin{cases} -x^2 + (a-c-y)x & \text{if } x+y \leq a, \\ -cx & \text{else.} \end{cases}
\end{aligned}
$$

Cournot competition

Profit:

- The **profit** for **Agent 1** per unit is therefore

$$
\begin{aligned}
\pi_1(x) &= rx - cx \\
&= \begin{cases} (a-(x+y))x - cx & \text{if } x+y \leq a, \\ 0 - cx & \text{else.} \end{cases} \\
&= \begin{cases} -x^2 + (a-c-y)x & \text{if } x+y \leq a, \\ -cx & \text{else.} \end{cases}
\end{aligned}
$$

Cournot competition

Profit:

- The **profit** for **Agent 1** per unit is therefore

$$\begin{aligned}
\pi_1(x) &= rx - cx \\
&= \begin{cases} (a-(x+y))x - cx & \text{if } x+y \leq a, \\ 0 - cx & \text{else.} \end{cases} \\
&= \begin{cases} -x^2 + (a-c-y)x & \text{if } x+y \leq a, \\ -cx & \text{else.} \end{cases}
\end{aligned}$$

- Likewise for **Agent 2**.

Cournot competition

Profit:

- The **profit** for **Agent 1** per unit is therefore

$$\begin{aligned}\pi_1(x) &= rx - cx \\ &= \begin{cases} (a-(x+y))x - cx & \text{if } x+y \leq a, \\ 0 - cx & \text{else.} \end{cases} \\ &= \begin{cases} -x^2 + (a-c-y)x & \text{if } x+y \leq a, \\ -cx & \text{else.} \end{cases}\end{aligned}$$

- Likewise for **Agent 2**.
- At (arbitrary or periodic) times the quantities x and y are revealed (simultaneously), and both agents adapt their production to the new situation (simultaneously).

Best response

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?
- Profit function is differentiable:

$$\frac{\partial}{\partial x}\pi_1(x) = \begin{cases} -2x + (a - c - y) & \text{if } x + y \leq a, \\ -c & \text{else.} \end{cases}$$

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?
- Profit function is differentiable:

$$\frac{\partial}{\partial x}\pi_1(x) = \begin{cases} -2x + (a - c - y) & \text{if } x + y \leq a, \\ -c & \text{else.} \end{cases}$$

- Stationary point:

$$x^* = \begin{cases} (a - c - y)/2 & \text{if } y \leq a - c, \\ 0 & \text{else.} \end{cases}$$

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?
- Profit function is differentiable:

$$\frac{\partial}{\partial x}\pi_1(x) = \begin{cases} -2x + (a - c - y) & \text{if } x + y \leq a, \\ -c & \text{else.} \end{cases}$$

- Stationary point:

$$x^* = \begin{cases} (a - c - y)/2 & \text{if } y \leq a - c, \\ 0 & \text{else.} \end{cases}$$

- 2nd-order derivative at x^* is positive

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?
- Profit function is differentiable:

$$\frac{\partial}{\partial x}\pi_1(x) = \begin{cases} -2x + (a - c - y) & \text{if } x + y \leq a, \\ -c & \text{else.} \end{cases}$$

- Stationary point:

$$x^* = \begin{cases} (a - c - y)/2 & \text{if } y \leq a - c, \\ 0 & \text{else.} \end{cases}$$

- 2nd-order derivative at x^* is positive $\Rightarrow$ π_1 is concave at x^*

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?
- Profit function is differentiable:

$$\frac{\partial}{\partial x}\pi_1(x) = \begin{cases} -2x + (a - c - y) & \text{if } x + y \leq a, \\ -c & \text{else.} \end{cases}$$

- Stationary point:

$$x^* = \begin{cases} (a - c - y)/2 & \text{if } y \leq a - c, \\ 0 & \text{else.} \end{cases}$$

- 2nd-order derivative at x^* is positive $\Rightarrow$ π_1 is concave at x^* $\Rightarrow$ maximum

Best response

- Suppose **Agent 1** knows the production quantity, y, of **Agent 2**. What would be the best response?
- Profit function is differentiable:

$$\frac{\partial}{\partial x}\pi_1(x) = \begin{cases} -2x + (a - c - y) & \text{if } x + y \leq a, \\ -c & \text{else.} \end{cases}$$

- Stationary point:

$$x^* = \begin{cases} (a - c - y)/2 & \text{if } y \leq a - c, \\ 0 & \text{else.} \end{cases}$$

- 2nd-order derivative at x^* is positive $\Rightarrow$ π_1 is concave at x^* $\Rightarrow$ maximum $\Rightarrow$ best response.

Outline

Cournot equilibrium

Cournot equilibrium

- If both production quantities are stationary, we speak of a **Cournot equilibrium**:

$$x^* = \begin{cases} (a - c - y^*)/2 & \text{if } y^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

Cournot equilibrium

- If both production quantities are stationary, we speak of a **Cournot equilibrium**:

$$x^* = \begin{cases} (a - c - y^*)/2 & \text{if } y^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

$$y^* = \begin{cases} (a - c - x^*)/2 & \text{if } x^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

Cournot equilibrium

- If both production quantities are stationary, we speak of a **Cournot equilibrium**:

$$x^* = \begin{cases} (a-c-y^*)/2 & \text{if } y^* \leq a-c, \\ 0 & \text{else.} \end{cases}$$
$$y^* = \begin{cases} (a-c-x^*)/2 & \text{if } x^* \leq a-c, \\ 0 & \text{else.} \end{cases}$$

Solving yields

Cournot equilibrium

- If both production quantities are stationary, we speak of a **Cournot equilibrium**:

$$x^* = \begin{cases} (a - c - y^*)/2 & \text{if } y^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

$$y^* = \begin{cases} (a - c - x^*)/2 & \text{if } x^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

Solving yields

$$(x^*, y^*) = \left(\frac{a-c}{3}, \frac{a-c}{3} \right).$$

Cournot equilibrium

- If both production quantities are stationary, we speak of a **Cournot equilibrium**:

$$x^* = \begin{cases} (a - c - y^*)/2 & \text{if } y^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

$$y^* = \begin{cases} (a - c - x^*)/2 & \text{if } x^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

Solving yields

$$(x^*, y^*) = \left(\frac{a - c}{3}, \frac{a - c}{3} \right).$$

Cournot equilibrium

- If both production quantities are stationary, we speak of a **Cournot equilibrium**:

$$x^* = \begin{cases} (a - c - y^*)/2 & \text{if } y^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$
$$y^* = \begin{cases} (a - c - x^*)/2 & \text{if } x^* \leq a - c, \\ 0 & \text{else.} \end{cases}$$

  Solving yields

$$(x^*, y^*) = \left(\frac{a - c}{3}, \frac{a - c}{3} \right).$$

- The Cournot equilibrium is not necessarily Pareto dominant.

Cournot dynamics

Possible questions:

Cournot dynamics

Possible questions:

1. How does adaptation proceed?

Cournot dynamics

Possible questions:

1. How does adaptation proceed?
2. Does it converge?

Cournot dynamics

Possible questions:

1. How does adaptation proceed?
2. Does it converge?
3. If yes, does every begin situation (x, y) leads to the same outcome?

Cournot dynamics

Possible questions:

1. How does adaptation proceed?
2. Does it converge?
3. If yes, does every begin situation (x, y) leads to the same outcome?
4. Are these outcomes (static) equilibria?

Cournot dynamics

Possible questions:

1. How does adaptation proceed?
2. Does it converge?
3. If yes, does every begin situation (x, y) leads to the same outcome?
4. Are these outcomes (static) equilibria?
5. Is it possible have more than one equilibrium?

Cournot dynamics

Possible questions:

1. How does adaptation proceed?
2. Does it converge?
3. If yes, does every begin situation (x, y) leads to the same outcome?
4. Are these outcomes (static) equilibria?
5. Is it possible have more than one equilibrium?
6. If yes, is it possible that some equilibria are missed by adaptation / iteration?

Outline

Outline

1. Introduction
   - Multi-agent learning (MAL)
   - Teaching
2. Cournot dynamics
   - Cournot competition
   - Cournot equilibrium
3. Alternative Cournot dynamics
   - Work of David Rand (1978)
   - Work of Tönu Puu (1991)
   - Conclusions

“Exotic Phenomena in Games and Duopoly”

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

Journal of Mathematical Economics 5 (1978) 173–184. © North-Holland Publishing Company

EXOTIC PHENOMENA IN GAMES AND DUOPOLY MODELS

David RAND

Mathematics Institute, University of Warwick, Coventry, U.K.

Received March 1977, final version received November 1977

1. Introduction

In this short note we draw attention to some very complex behaviour which occurs in very simple games and, in particular, in Cournot duopoly and its generalisations [Cournot (1834) and Wald (1951)].

“Exotic Phenomena in Games and Duopoly”

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

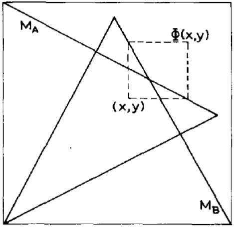


Fig. 1

It will be clear that the results which follow will not be greatly affected if we replace the curves by any segments which approximate them provided the slope is always greater than some constant greater than one. Any difficulties about the structural stability of the results arises from the non-differentiability of the curves at $z_A = (1-\varepsilon_1, \frac{1}{2})$ and $z_B = (\frac{1}{2}, 1-\varepsilon_2)$. Otherwise the results are perfectly stable.

"Exotic Phenomena in Games and Duopoly"

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

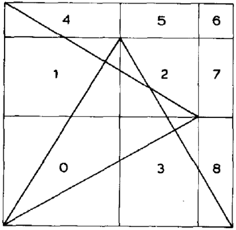


Fig. 2

Thus $[i_0 i_1 \dots i_n] = [i_0] \cap \Phi^{-1}(i_1 i_2 \dots i_n])$. Clearly it is possible that $[i_0 i_1 \dots i_n] = \emptyset$.

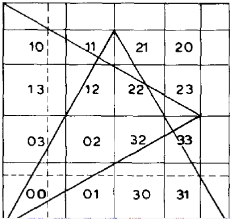
Equivalently the rectangles can be constructed by drawing in the vertical lines through the two points of intersection of l and M_B and the two points of intersection of h and M_A. This cuts

“Exotic Phenomena in Games and Duopoly”

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

sets. Moreover, since ε_1 and ε_2 are chosen very small there exists a constant F satisfying $0<F<1$ such that for any such sequence $|i_0 i_1 \dots i_n| < F|i_1 i_2 \dots i_n|$ where $|i|$ denotes the maximum of the height and width of $[i]$. Thus, by induction $|i_0 i_1 \dots i_n| < F^n |i_n| < F^n$, which proves that $|i_0 i_1 \dots i_n| \to 0$ as $n \to +\infty$, whence $[i]$ consists of a single point.

We assume now that the situation is as in fig. 3, so that $\Phi(I^2 \backslash J)$ is contained in the union of the rectangles [10], [13], [03], [00], [01], [30] and [31].



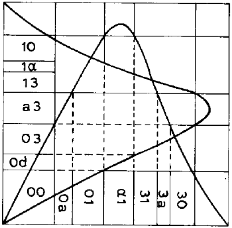
“Exotic Phenomena in Games and Duopoly”

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

182 *D. Rand, Games and duopoly models*

where $i_0, i_1, \ldots, i_n \in \Omega$.

To simplify the exposition we shall also assume that h meets M_A in the interior of $[10] \cup [40]$ and l meets M_B in the interior of $[30] \cup [80]$. Let $\hat{\Omega}$ denote the set {01,1d,13,a3,03,0d,00,0a,01,d1,31,3a,30}. Then this just says that if x belongs to the union of [4], [5], [6], [7], [8], [α] and [β] then $\Phi(x) \in [ij]$ where $ij \in \hat{\Omega}$ (see fig. 6).

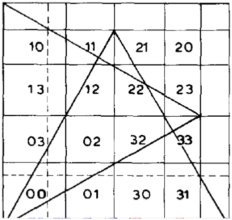


"Exotic Phenomena in Games and Duopoly"

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

sets. Moreover, since ε_1 and ε_2 are chosen very small there exists a constant F satisfying $0<F<1$ such that for any such sequence $|i_0 i_1 \dots i_n| < F|i_1 i_2 \dots i_n|$ where $|i|$ denotes the maximum of the height and width of $[i]$. Thus, by induction $|i_0 i_1 \dots i_n| < F^n |i_n| < F^n$, which proves that $|i_0 i_1 \dots i_n| \to 0$ as $n \to +\infty$, whence $[i]$ consists of a single point.

We assume now that the situation is as in fig. 3, so that $\Phi(I^2 \backslash J)$ is contained in the union of the rectangles [10], [13], [03], [00], [01], [30] and [31].



“Exotic Phenomena in Games and Duopoly”

David Rand. Journal of Math. Ec., 1978, vol. 5(2), pp. 173-184.

instead of the two lines $y=C_A$, $x=C_B$ as in Example 1 to cope with the fact that the components of the preimages of rectangles intersecting this area may be larger in size than the original rectangle. This prevents us from obtaining a direct analogue of Lemma 3.2. Now continue this decomposition by constructing the subsets

$$[i_0 i_1 \dots i_n] = \{x \in I^2 | \Phi^j(x) \in [i_j] \text{ for } 0 \leqq j \leqq n\},$$

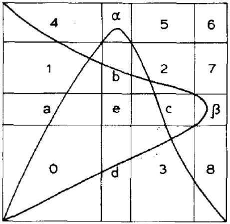


Fig. 5

Outline

1. Introduction
   - Multi-agent learning (MAL)
   - Teaching
2. Cournot dynamics
   - Cournot competition
   - Cournot equilibrium
3. Alternative Cournot dynamics
   - Work of David Rand (1978)
   - Work of Tönu Puu (1991)
   - Conclusions

Alternative assumptions

Proposed around 1991

Alternative assumptions

Proposed around 1991 by

- *Tönu Puu* (born 1936, in Tallinn, Estonia).

Alternative assumptions

Proposed around 1991 by

- *Tönu Puu* (born 1936, in Tallinn, Estonia).
- Professor of Economics Emeritus at Umeå University, Sweden.
- Portfolio selection, investment and production, phil. of science, spatial economics, nonlinear dynamic processes, oligopoly, business cycles.

Alternative assumptions

Proposed around 1991 by

- *Tönu Puu* (born 1936, in Tallinn, Estonia).
- Professor of Economics Emeritus at Umeå University, Sweden.
- Portfolio selection, investment and production, phil. of science, spatial economics, nonlinear dynamic processes, oligopoly, business cycles.

Idea: suppose

Alternative assumptions

Proposed around 1991 by

- *Tönu Puu* (born 1936, in Tallinn, Estonia).
- Professor of Economics Emeritus at Umeå University, Sweden.
- Portfolio selection, investment and production, phil. of science, spatial economics, nonlinear dynamic processes, oligopoly, business cycles.

Idea: suppose

- **Sales price** per unit is iso-elastic:

$$s =_{Def} \max\left\{0, \frac{1}{x+y}\right\}.$$

Alternative assumptions

Proposed around 1991 by

- *Tönu Puu* (born 1936, in Tallinn, Estonia).
- Professor of Economics Emeritus at Umeå University, Sweden.
- Portfolio selection, investment and production, phil. of science, spatial economics, nonlinear dynamic processes, oligopoly, business cycles.

Idea: suppose

- **Sales price** per unit is iso-elastic:

$$s =_{Def} \max \left\{ 0, \frac{1}{x+y} \right\}.$$

- Agents have different **production costs** α and β per unit.

Alternative assumptions

Proposed around 1991 by

- *Tönu Puu* (born 1936, in Tallinn, Estonia).
- Professor of Economics Emeritus at Umeå University, Sweden.
- Portfolio selection, investment and production, phil. of science, spatial economics, nonlinear dynamic processes, oligopoly, business cycles.

Idea: suppose

- **Sales price** per unit is iso-elastic:

$$s =_{Def} \max\left\{0, \frac{1}{x+y}\right\}.$$

- Agents have different **production costs** α and β per unit.
- **Adaptation** proceeds gradually, through learning:

$$\text{new} = (1-\delta)\cdot \text{old} + \delta \cdot \text{input}.$$

New response functions

- Compute response functions as usual.

New response functions

- Compute response functions as usual. We arrive at:

New response functions

- Compute response functions as usual. We arrive at:

$$x^* = \sqrt{\frac{y}{\alpha}} - y$$

New response functions

- Compute response functions as usual. We arrive at:

$$x^* = \sqrt{\frac{y}{\alpha}} - y$$

$$y^* = \sqrt{\frac{x}{\beta}} - x$$

New response functions

- Compute response functions as usual. We arrive at:

$$\begin{cases} x^* & = \sqrt{\dfrac{y}{\alpha}} - y \\[2ex] y^* & = \sqrt{\dfrac{x}{\beta}} - x \end{cases}$$

New response functions

- Compute response functions as usual. We arrive at:

$$\begin{cases} x^* &= \sqrt{\dfrac{y}{\alpha}} - y \\ \\ y^* &= \sqrt{\dfrac{x}{\beta}} - x \end{cases}$$

- Cournot: adapt simultaneously.

New response functions

- Compute response functions as usual. We arrive at:

$$\begin{cases} x^* &= \sqrt{\dfrac{y}{\alpha}} - y \\ \\ y^* &= \sqrt{\dfrac{x}{\beta}} - x \end{cases}$$

- Cournot: adapt simultaneously.

$$x_{t+1} = \sqrt{\frac{y_t}{\alpha}} - y_t$$

New response functions

- Compute response functions as usual. We arrive at:

$$\begin{cases} x^* &= \sqrt{\dfrac{y}{\alpha}} - y \\ \\ y^* &= \sqrt{\dfrac{x}{\beta}} - x \end{cases}$$

- Cournot: adapt simultaneously.

$$\begin{aligned} x_{t+1} &= \sqrt{\frac{y_t}{\alpha}} - y_t \\ y_{t+1} &= \sqrt{\frac{x_t}{\beta}} - x_t \end{aligned}$$

New response functions

- Compute response functions as usual. We arrive at:

$$\begin{cases} x^* & = \sqrt{\dfrac{y}{\alpha}} - y \\ y^* & = \sqrt{\dfrac{x}{\beta}} - x \end{cases}$$

- Cournot: adapt simultaneously.

$$\begin{cases} x_{t+1} & = \sqrt{\dfrac{y_t}{\alpha}} - y_t \\ y_{t+1} & = \sqrt{\dfrac{x_t}{\beta}} - x_t \end{cases}$$

New response functions

- Puu: adapt gradually.

New response functions

- Puu: adapt gradually.

$$x_{t+1} = x_t + \theta \left(\sqrt{\frac{y_t}{\alpha}} - y_t - x_t \right)$$

New response functions

- Puu: adapt gradually.

$$x_{t+1} = x_t + \theta \left(\sqrt{\frac{y_t}{\alpha}} - y_t - x_t \right)$$

where $0 \leq \theta \leq 1$ represents the **learning speed**.

New response functions

- Puu: adapt gradually.

$$x_{t+1} = x_t + \theta \left(\sqrt{\frac{y_t}{\alpha}} - y_t - x_t \right)$$

$$y_{t+1} = y_t + \theta \left(\sqrt{\frac{x_t}{\beta}} - x_t - y_t \right),$$

where $0 \leq \theta \leq 1$ represents the **learning speed**.

New response functions

- Puu: adapt gradually.

$$\begin{cases} x_{t+1} & = x_t + \theta \left(\sqrt{\frac{y_t}{\alpha}} - y_t - x_t \right) \\ y_{t+1} & = y_t + \theta \left(\sqrt{\frac{x_t}{\beta}} - x_t - y_t \right), \end{cases}$$

where $0 \leq \theta \leq 1$ represents the **learning speed**.

New response functions

- Puu: adapt gradually.

$$\begin{cases} x_{t+1} & = x_t + \theta \left(\sqrt{\dfrac{y_t}{\alpha}} - y_t - x_t \right) \\ y_{t+1} & = y_t + \theta \left(\sqrt{\dfrac{x_t}{\beta}} - x_t - y_t \right), \end{cases}$$

  where $0 \leq \theta \leq 1$ represents the **learning speed**.
- If $\alpha/\beta \in (3 - 2\sqrt{2}, 3 + 2\sqrt{2})$ then Nash equilibrium is stable.

New response functions

- Puu: adapt gradually.

$$\begin{cases} x_{t+1} & = x_t + \theta \left(\sqrt{\frac{y_t}{\alpha}} - y_t - x_t \right) \\ y_{t+1} & = y_t + \theta \left(\sqrt{\frac{x_t}{\beta}} - x_t - y_t \right), \end{cases}$$

  where $0 \leq \theta \leq 1$ represents the **learning speed**.
- If $\alpha/\beta \in (3 - 2\sqrt{2}, 3 + 2\sqrt{2})$ then Nash equilibrium is stable.
- If $\alpha/\beta \in [4/25, 25/4]$ then trajectory remains bounded.

New response functions

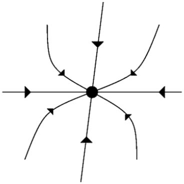
- Puu: adapt gradually.

$$\begin{cases} x_{t+1} & = x_t + \theta \left(\sqrt{\frac{y_t}{\alpha}} - y_t - x_t \right) \\ y_{t+1} & = y_t + \theta \left(\sqrt{\frac{x_t}{\beta}} - x_t - y_t \right), \end{cases}$$

  where $0 \leq \theta \leq 1$ represents the **learning speed**.
- If $\alpha/\beta \in (3 - 2\sqrt{2}, 3 + 2\sqrt{2})$ then Nash equilibrium is stable.
- If $\alpha/\beta \in [4/25, 25/4]$ then trajectory remains bounded.
- If $0.16 \leq \alpha/\beta \leq 0.171$ or $5.828 \leq \alpha/\beta \leq 6.25$ then bounded but not stable $\Rightarrow$ periodicity, semi-periodicity, or chaos.

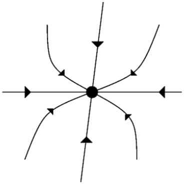
Four types of system behaviour

Four types of system behaviour

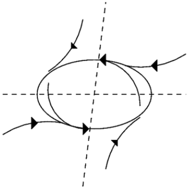


Stable

Four types of system behaviour

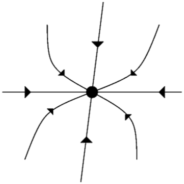


Stable

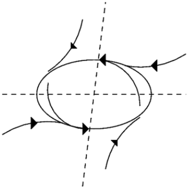


Periodic

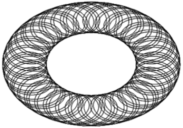
Four types of system behaviour



Stable

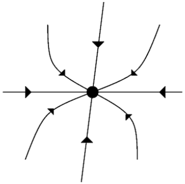


Periodic

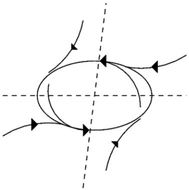


Quasiperiodic

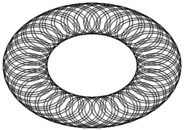
Four types of system behaviour



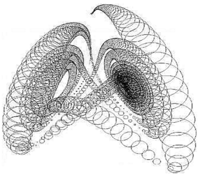
Stable



Periodic



Quasiperiodic



Chaotic

Outline

1. Introduction
   - Multi-agent learning (MAL)
   - Teaching
2. Cournot dynamics
   - Cournot competition
   - Cournot equilibrium
3. Alternative Cournot dynamics
   - Work of David Rand (1978)
   - Work of Tönu Puu (1991)
   - Conclusions

Conclusions

Conclusions:

Conclusions

Conclusions:

- Multi-agent learning is not only about **learning**. It is also about teaching.

Conclusions

Conclusions:

- Multi-agent learning is not only about **learning**. It is also about teaching.
- **Simple scenarios** of multi-agent learning may invoke complex behaviour.

Conclusions

Conclusions:

- Multi-agent learning is not only about **learning**. It is also about teaching.
- **Simple scenarios** of multi-agent learning may invoke complex behaviour.