INFO-F-409 Learning dynamics

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

T. Lenaerts and Y.-M. De Hauwere MLG, Université Libre de Bruxelles and Al-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 I

The formation of agents' beliefs

Now that we can determine the Nash and subgame 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 :

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

Levels of learning

innate

reflex actions





Conditioning



Observational and imitative learning

Teaching









© Tom Lenaerts, 2010

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 action the other players' actions in the previous round

An action profile that remains the same from period to period is a pure Nash equilibrium of the game

Best-response dynamics



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			BELIEF				BELIEF		
	A plays	B plays		A plays	B plays		A plays	B plays	
prior	В	В	prior	S	S	prior	S	В	
I	В	В	I	S	S	I.	В	S	
2	В	В	2	S	S	2	S	В	
•••	•••	•••	•••	•••	•••	•••	•••	•••	

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 I

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:

		BELIEF			
		A plays		B plays	
prior		(1,0)		(0,1)	
I	S	(I,I)	В	(I,I)	TOTAL = 2
2	S	(1,2)	S	(I,2)	TOTAL = 3
3	S	(2,2)	S	(1,3)	TOTAL = 4
4	S	(2,3)	В	(2,3)	TOTAL = 5
5	S	(2,4)	S	(2,4)	TOTAL = 6
6	S	(2,5)	S	(2,5)	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



observation

More details on reinforcement learning by prof. De Hauwere

Culture and evolution

"Culture is the integrated pattern of human knowledge, belief, and behavior that depends upon the capacity for learning and transmitting knowledge to succeeding generations"

Link between biological evolution and cultural learning

A trait is adaptive if **biology** \rightarrow it provides an increase in an individual's chance of survival and reproductive success **culture** \rightarrow it provides an advantage in the interactions with other players



© Tom Lenaerts, 2010

Biological analog

This transmission can imply copying (reproduction), with the possibility of errors (mutation)





evolutionary dynamics

The thing that evolves is the population



Individuals reproduce in the population



Individuals imitate other players in the population

Individuals may mutate during the reproductive process



Individuals may make errors when imitating

Selection may cause the new individuals to replace the others



Selection may cause certain behaviors to be imitated more often than the others

Selection is the outcome of a competition between the different types of population members



Social dilemmas

THE QUESTION OF COOPERATION

Social dilem a group has a choice thatpoorer outco if none had a individually

Social dilemmas are situations in which individual rationality leads to collective irrationality. That is, individually reasonable behavior leads to a situation in which everyone is worse off than they might have been otherwise. Many of the most challenging programs we face. from the interpersonal to the interna-

P. Kollock (1998) Social Dilemmas: the anatomy of cooperation Ann. Rev. Sociol. 24:183-214

doing less well than they would have done if they had acted unreasonably or irrationally. This paradoxical pos-

R.M. Dawes and D.M. Messick (2000) Social Dilemmas. International Journal of Psychology 35(2):111-116

The Tragedy of the Commons

The population problem has no technical solution; it requires a fundamental extension in morality.

Garrett Hardin

At the end of a thoughtful article on future of nuclear war, Wiesner and 'k (1) concluded that: "Both sides in Irms race are ... confronted by the ama of steadily increasing military er and steadily decreasing national rity. It is our considered profesal judgment that this dilemma has technical so

continue area of result ** would t on the nal secu the kind nely tha a to the lost uni ns publ nipopula proble inical se y be de inge onl

sional judgment. . . ." Whether they were right or not is not the concern of the present article. Rather, the concern here is with the important concept of a class of human problems which can be called "no technical solution problems," and, more specifically, with the identification and discussion of one of these.

ow that the class is not If the game of tickhe problem, "How of tick-tack-toe?" t I cannot, if I asth the conventions t my opponent unperfectly. Put anno "technical solun. I can win only eaning to the word opponent over the im; or I can falsify iy in which I "win" ense, an abandon-

we intuitively unalso, of course, game—refuse to Population, as Malthus said, natura tends to grow "geometrically," or, as would now say, exponentially. In finite world this means that the capita share of the world's goods n steadily decrease. Is ours a finite wo

A fair defense can be put forward the view that the world is infinite that we do not know that it is not. in terms of the practical problems we must face in the next few ger tions with the foreseeable technolog is clear that we will greatly incr human misery if we do not, during immediate future, assume that the w available to the terrestrial human p ulation is finite. "Space" is no esc (2).

A finite world can support only finite population; therefore, popula growth must eventually equal zero. (I case of perpetual wide fluctuati above and below zero is a trivial vari that need not be discussed.) When condition is met, what will be the sit tion of mankind? Specifically, can B tham's goal of "the greatest good the greatest number" be realized?

No—for two reasons, each sufficient by itself. The first is a theoretical or It is not mathematically possible maximize for two (or more) variables the same time. This was clearly stat by von Neumann and Morgenstern (. but the principle is implicit in the theoret of partial differential equations, dati back at least to D'Alembert (171) 1783).

donrunfrom biological facts. To live, a urse, organism must have a source of ener e to (for the food). This energy

What?

G. Hardin (1968) The tragedy of the commons. Science 168:1243-1248

"commons" originated in medieval England, a piece of land to which people had access for free

Discusses the disastrous effects that individual selfish choice may have on common resources and global welfare.

Highlights the issue of unlimited population growth and the limited size of our world

G. Hardin (1915-2003)

<u>Assumption I</u>; a number of herdsman Nwho have access to a commons, which is used for grazing by their cattle. Assumption 2; Each herdsman is expected to keep as many cattle as possible as this provides him with profit.

What is the utility **to one herdsman** of adding one more animal to his herd? <u>Positive</u>: he receives additional profit from the sale (+1)

<u>Negative</u>: additional grazing, for which the cost is shared with the other herdsman (-1/N)



<u>Assumption I</u>; a number of herdsman Nwho have access to a commons, which is used for grazing by their cattle. <u>Assumption 2</u>; Each herdsman is expected to keep as many cattle as possible as this provides him with profit.

What is the utility **to one herdsman** of adding one more animal to his herd?

Since the benefit outweighs (+1) the cost (-1/N); add another animal



<u>Assumption I</u>; a number of herdsman Nwho have access to a commons, which is used for grazing by their cattle.

When all *N* herdsman reach the same conclusion



<u>Assumption 2</u>; Each herdsman is expected to keep as many cattle as possible as this provides him with profit.

"Ruin is the destination towards which all men rush, each pursuing his own best interest in a society that beliefs in the freedom of the commons" (G. Hardin, 1968)











Our commons







cooperation

noun

- 1 cooperation between management and workers: COLLABORATION, joint action, combined effort, teamwork, partnership, coordination, liaison, association, synergy, synergism, give and take, compromise.
- 2 thank you for your cooperation: ASSISTANCE, helpfulness, help, helping hand, aid.





Cooperation?





pays a cost c

receives a benefit b > c



© Tom Lenaerts, 2013

Cooperation?



pays a cost c



receives a benefit b > c



Cooperation?

pays a cost cImage: Colspan="2">Image: Colspan="2"/Image: Colspan=""/Image: Colsp

receives a benefit b > cIt's better to play D, when the opponent plays C

It's better to play D, when the opponent plays D

But CC is better than DD



Fear AND Greed

T = b > R = b - c > P = 0 > S = -c



greed = T > R

fear = P > S

R = reward S = suckers payoff T = temptation to defect P = punishment

C.H. Coombs (1973) A reparameterization of the prisoner's dilemma game. Behavioral Science 18:424-428



Fear AND Greed D 3 ()

greed = T > R

fear = P > S

Best response determines

7

D



Nash equilibrium

C.H. Coombs (1973) A reparameterization of the prisoner's dilemma game. Behavioral Science 18:424-428

Fear OR Greed



no greed, only fear





© Tom Lenaerts, 2013

D

7



no fear, only greed

Snow-drift game

Evolutionary stable strategies



Can a C player invade a population of D players?

The fraction of C (D) players is ε (1- ε)

 $S(1-\varepsilon)+R\varepsilon > P(1-\varepsilon)+T\varepsilon$

C can invade when:

i) S > P or ii) S = P and R > T



Can D imade D?



















Evolutionary dynamics



Replicator equation ... $\frac{dx}{dt} = x(1-x)[fc(x)-fb(x)]$ = x(1-x)[(b-c+c-b+0)x-c-0] = -cx(1-x)

Evolutionary Games and Population Dynamics

sef Hofbauer and Karl Sigmund

OLUTIONARY

DYNAMICS

MARTIN A.

P.D. Taylor and L.B. Jonker (1978) Evolutionary stable strategies and game dynamics. Mathematical biosciences 40(1-2):145-156
Dynamics of social dilemmas

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}(1-\mathbf{x})[(R-S-T+P)\mathbf{x}+S-P]$$



Equilibria Summary



How to evolve cooperation





Indirect reciprocity



M. Nowak (2006) Five rules for the evolution of cooperation. Science 314:1560-1563

kin selection



group selection



C.Taylor and M. Nowak (2007) Transforming the dilemma. Evolution 61-10:2281-2292

How would the dynamics change when interactions between the same two individuals can be **repeated**?



What kind of strategies could we make that take into account the actions from previous encounters?

R.Trivers(1971) The evolution of reciprocal altruism Q Rev Biol 46:35-37

In 1978 R. Axelrod organised a tournament between strategies that play against each other in a Prisoner's dilemma (14 contestants)

TFT - Tit For Tat (A. Rapoport)

For cooperation to emerge :

I) individuals should be involved in **ongoing relationship**

- 2) individuals should be able to identify each other
- 3) posses information about how individuals behaved in the

past→enormous strategy space



Thus TFT plays C in the first round and then plays the same strategy as the opponents previous one

With probability w there is another round of interactions On average there are 1/(1-w) interactions between two players







Indirect reciprocity takes your **reputation** into account



"Indirect reciprocity arises out of direct reciprocity in the presence of an interested audience"

> R.Alexander (1987) The Biology of Moral Systems. Aldine de Gruyter, NY

Interactions are not repeated in this context



Each individual has a reputation which initialised to 0

Every time the individual helps someone, the reputation increases

Whenever the individual refuses to help, the reputation decreases

Conditional strategies could take advantage of this information: Help is only provided to another individual when the reputation exceeds a certain value

Assume a strategy DISC that will cooperate unless it knows that the other strategy is a defector

The probability of knowing the strategy is q, so (1-q) is the probability that DISC will cooperate with a defector



DISC is an ESS when R > (1-q)T + qP

$$q > \frac{T - R}{T - P}$$

Take for instance the values for the prisoner's dilemma



Small summary

So in case of direct and indirect reciprocity one obtains very similar results in the prisonners dilemma



The prisoner's dilemma is transformed into a stag-hunt game !





F.C. Santos, J.M. Pacheco and T. Lenaerts (2006) Evolutionary dynamics in social dilemmas in structured heterogeneous populations. Proc Natl Acad Sci USA 103:3490-3494

Evolutionary dynamics

 $k_3 = 2$ C $k_l = l$ $k_2 = 4$ $k_4 = 5$ D $k_5=2$ D $k_6=2$ $k_{10} = 1$ D $k_8 = 5$ С D D $k_7 = 3$

Stochastic replicator equation ...

Vertex x plays k_x times and accumulates payoff f_x

Choose a random neighbor y (payoff f_y)

Replace strategy S_x by S_y with probability $p=max[0, (f_y-f_x)/k_>(T-S)]$

Simulation I



100 runs 50% **C**, 50% **D** *R*=1, *P*=0



Which networks?



Which models have people been using?

What does data tell us about real networks?



Regular graphs

Every node has exactly the same degree < k > = 4

regular and democratic network



Simulation II





Random graphs



p=1/6, N=10 $\Rightarrow < k > \approx 1.5$



P. Erdős (1913-1996)

random and democratic network

Simulation III

There is a limit to the number of interactions but its democratic and random

 $N = 10^4 < k > = 4$

100 runs 50% C, 50% D *R*=1, *P*=0



Which networks?



Which models have people been using?

What does data tell us about real networks?



Small world experiment

What is the average number of connections between any two people?





S. Milgram (1933-1984)

"Six degrees of separation" (J. Guare, 1990)

J. Travers and S. Milgram (1969) An experimental study of the small-world problem. Sociometry 32(4):425-443



The average path length (L) is a measure of proximity between nodes

Cluster coefficient



The cluster coefficient (C) is a measure for cliquishness

Small world networks





Mechanism:

I.take a regular graph2.randomly rewire every edge with probability p3.avoid loops and double edges

Network classes

Classes of small-world networks

L. A. N. Amaral*, A. Scala, M. Barthélémy[†], and H. E. Stanley

Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215

Communicated by Herman Z. Cummins, City College of the City University of New York, New York, NY, July 13, 2000 (received for review April 20, 2000)

We study the statistical properties of a variety of diverse real-world networks. We present evidence of the occurrence of three classes of small-world networks: (a) scale-free networks, characterized by a vertex connectivity distribution that decays as a power law; (b) broad-scale networks, characterized by a connectivity distribution that has a power law regime followed by a sharp cutoff; and (c) single-scale networks, characterized by a connectivity distribution with a fast decaying tail. Moreover, we note for the classes of broad-scale and single-scale networks that there are constraints limiting the addition of new links. Our results suggest that the nature of such constraints may be the controlling factor for the emergence of ifferent charge of networks these networks, there are constraints limiting the addition of new links. Our results suggest that such constraints may be the controlling factor for the emergence of scale-free networks.

Empirical Results

First, we consider two examples of technological and economic networks: (*i*) the electric power grid of Southern California (2), the vertices being generators, transformers, and substations and the links being high-voltage transmission lines; and (*ii*) the network of world airports (24), the vertices being the airports and the links being nonstop connections. For the case of the airport metwork, we have access to data on number of passengers in

L.A.N.Amaral, A. Scala, M. Barthelemy and H.E. Stanley (2000) Classes of small-world networks. Proc Natl Acad Sci USA 97(21): 11149-11152

electrical power grid of South California, network of world airports, movie-actor network, acquintance network of mormons, friendship network of 417 Madison Junior High school students, the neuronal network of the worm C.elegans, the conformational space of a lattice polymer chain



Network classes

Aging of vertices as in the movie-actor network

cost of adding links or the limited capacity of vertices as in the airport network

L.A.N.Amaral, A. Scala, M. Barthelemy and H.E. Stanley (2000) Classes of small-world networks. Proc Natl Acad Sci USA 97(21): 11149-11152

Configuration model



M. Molloy and B. Reed (1995) A critical point for random graphs with a given degree sequence. Random Struct. Algorithms 6:161-180

Simulation IV



Scale-free Networks



A.-L. Barabási and R. Albert (1999) Emergence of Scaling in Random Networks. Science 286:509-512

Simulation V



Simulation VI



intuition

Defectors are victims of their own success ...



social versus individual learning




social versus individual learning



Stimulus-response learning ...

Vertex x plays once with a random neighbor y and both receive a payoff f_x (and f_y)

update strategy using the following model $i \in \{x, y\}$:

 $p_i(t+1) = p_i(t) + \lambda \beta_i(t) * (1-p_i(t))$ when *i* played C at time *t*

 $p_i(t+1) = p_i(t) - \lambda \beta_i(t) * p_i(t)$ when *i* played **D** at time *t*

social versus individual learning





Presentation next week!



Networks are dynamic



Networks are dynamic

Agent-based simulations

F.C. Santos, J.M. Pacheco and T. Lenaerts (2006) Cooperation previals when individuals adjust their social ties. PLoS Comp Biol 2(12):e178

S.Van Segbroeck, F.C. Santos, A. Nowé, J.M. Pacheco and T. Lenaerts (2008) The evolution of prompt reactions to adverse ties. BMC Evol Biol 8:287

Analytics and numerical approximations

J.M Pacheco, A. Traulsen and M. Nowak (2006) Coevolution of strategy and structure in complex networks with dynamical linking. Phys Rev Lett 97:258103

S.Van Segbroeck, F.C. Santos, T. Lenaerts and J.M. Pacheco (2009) Reacting differently to adverse ties promotes the evolution of cooperation. Phys Rev Lett 102:058105



two timescales



Fast linking promotes C

Simulation



Effects on topology





Effects on topology





 $N=10^{4}$ <k>=30 $\beta=0.005$ 50% C, 50% D T=2, R=1,P=0, S=1



Everyone reacts differently





ROBERT AUDI





Change introduces uncertainty!

Evolution of rewiring



Fast versus slow

Simulation I



Recent experiments



Fehl, K., Van Der Post, D. J., & Semmann, D. (2011). Co-evolution of behaviour and social network structure promotes human cooperation. Ecology Letters, 14(6), 546–551.

© Tom Lenaerts, 2010

Summary

