MONASH University
Information Technology

FIT5226
Multi-agent System & Collective Behaviour

Wk 8: Population Games and Evolutionary Game Theory

Bernd Meyer, April 2025

Recall: The Price of Anarchy

The “Price of Anarchy” is the payoff lost due to the lack of central coordination.

For example, in the prisoners’ dilemma (below),
central coordination would enable the prisoners to achieve -2 in total
due to a lack of coordination they will only obtain -6 at Nash.

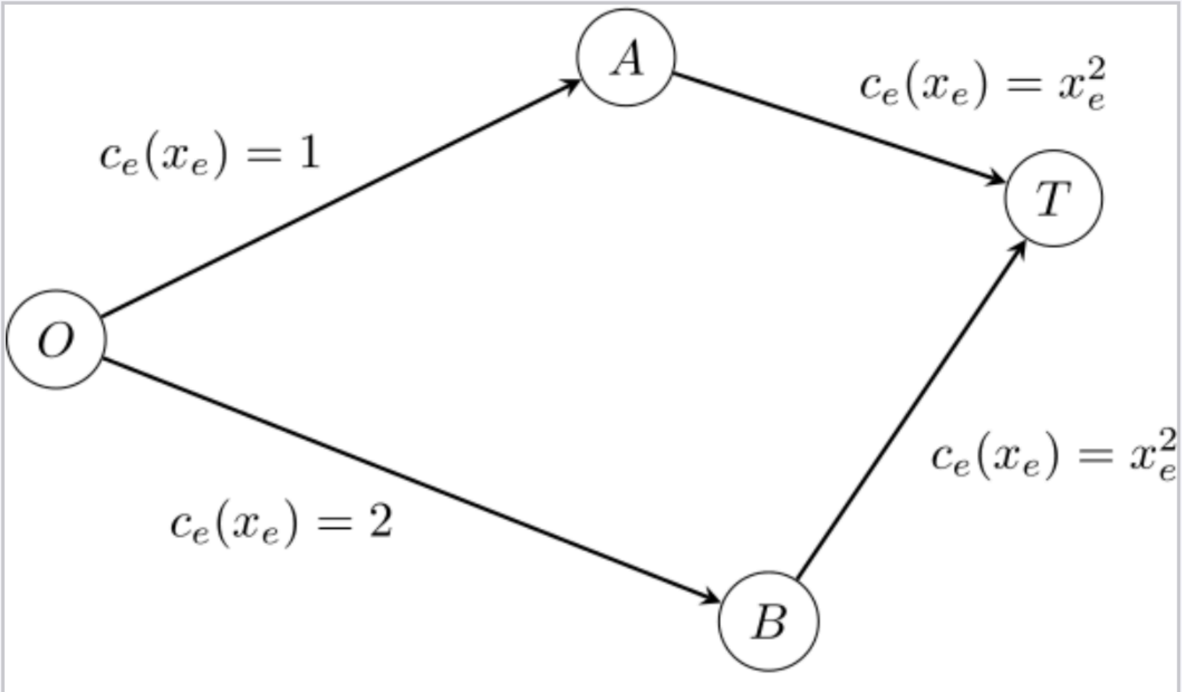
This “price of anarchy” is usually defined as the quotient of these quantities, here 2/6

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Example: Selfish Routing

Cost Matrix

p1 \ p2	A	B
A	(5,5)	(2,3)
B	(3,2)	(6,6)



The directed graph for a simple congestion game.

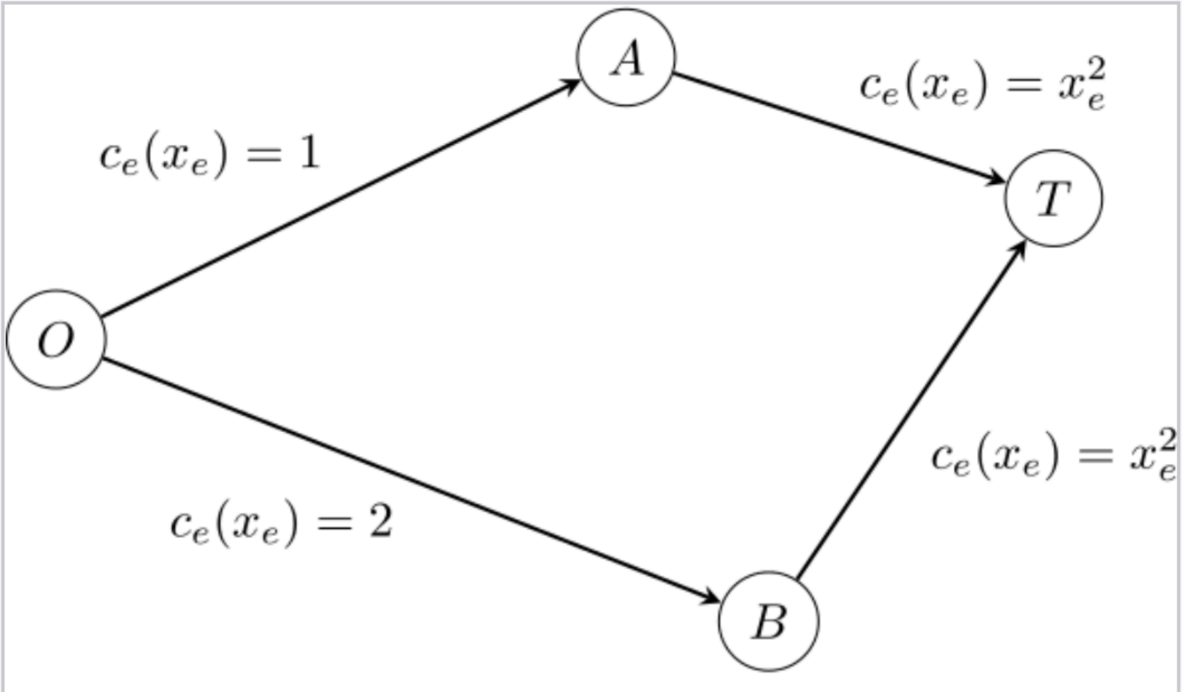
Nash=???

https://en.wikipedia.org/wiki/Congestion_game

Intro: Selfish Routing (Example)

Cost Matrix

p1 \ p2	A	B
A	(5,5)	(2,3)
B	(3,2)	(6,6)



The directed graph for a simple congestion game.

Nash=(A,B); (B,A)

https://en.wikipedia.org/wiki/Congestion_game

Congestion Games (Def)

Selfish routing is a particular form of a congestion game

- A base set of congestible elements E;
- n players;
- A finite set of strategies S_i for each player, where each strategy $P \in S_i$ is a subset of E;
- For each element e and a vector of strategies $(P_1, P_2, \ldots, P_n)$, a load $x_e = \#\{i : e \in P_i\}$;
- For each element e, a delay function $d_e : \mathbb{N} \longrightarrow \mathbb{R}$;
- Given a strategy P_i, player i experiences delay $\sum_{e \in P_i} d_e(x_e)$. Assume that each d_e is positive and monotone increasing.

https://en.wikipedia.org/wiki/Congestion_game

Population Games

Population games are played in large populations.

All **players are indistinguishable**.
They have the same strategy sets.
It does not matter, who plays which strategy.

Only the proportion of the population playing a particular strategy matters.

Observation: Congestion games are a particular form of population games.

Potential Games

A game $G = (N, A, u)$ is a potential game if there exists a (single) function $P : A \mapsto \mathbb{R}$ such that **for all** $i \in \mathbb{N}$, all $a_{-i} \in A\_i$ and a_i, a_i':

$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = P(a_i, a_{-i}) - P(a_i', a_{-i})$$

Theorem: Every congestion game is a potential game

Theorem: Every (finite) potential game has a <u>pure</u> Nash Equilibrium

Myopic Best-Response

always terminates with Nash for potential games
(but can cycle for general games)

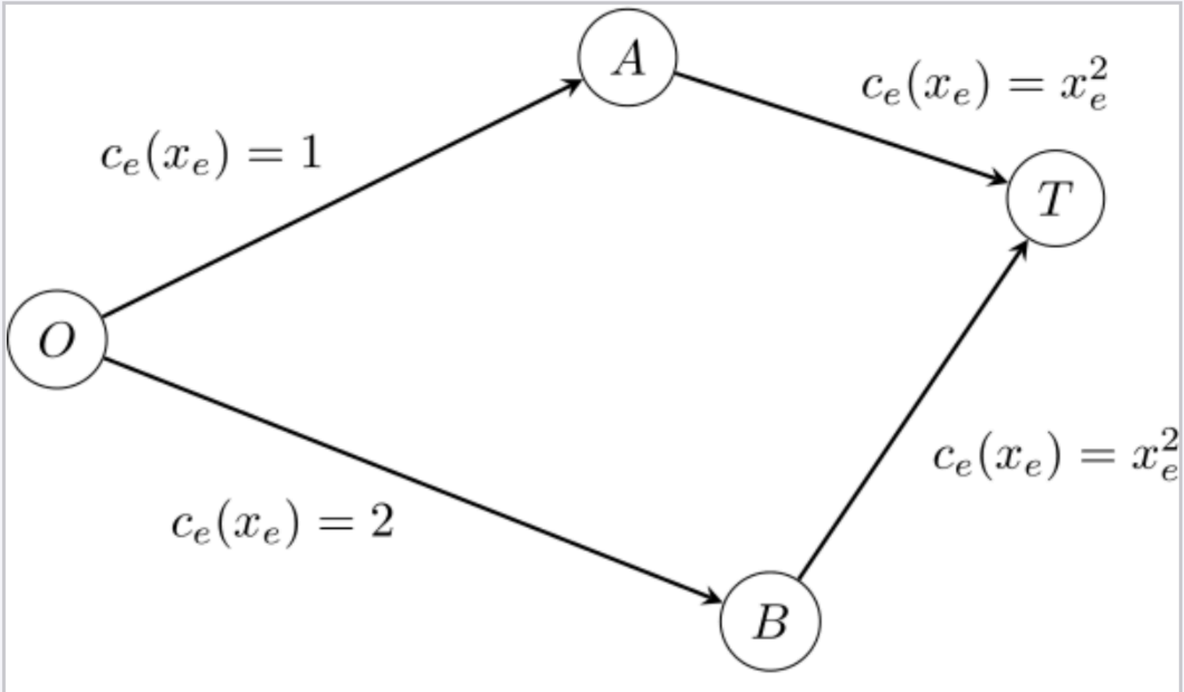
function MyopicBestResponse (game G, action profile a) **returns** a
while *there exists an agent i for whom a_i is not a best response to a_{-i}* **do**
  $a_i' \leftarrow$ some best response by i to a_{-i}
  $a \leftarrow (a_i', a_{-i})$
return a

but is this Nash Equilibrium necessarily pro-social?

Selfish Routing (Price of Anarchy)

Cost Matrix

p1 \ p2	A	B
A	(5,5)	(2,3)
B	(3,2)	(6,6)



The directed graph for a simple congestion game.

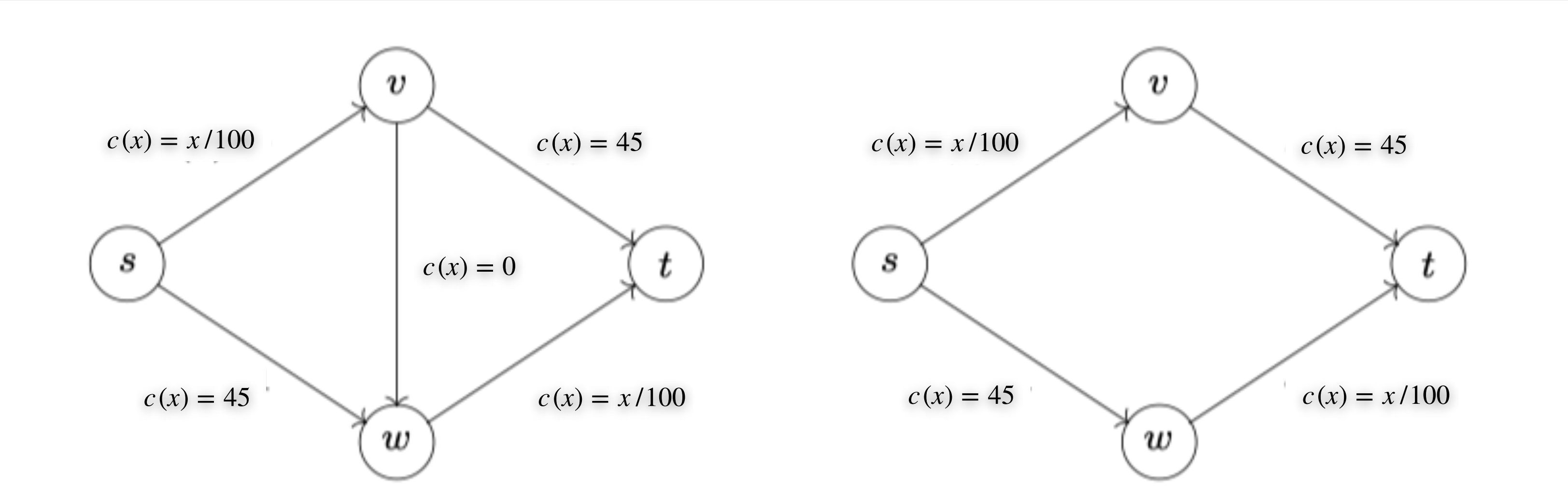
Nash=(A,B); (B,A)

https://en.wikipedia.org/wiki/Congestion_game

here, Nash is pro-social

Braess' Paradox *or* The problem with Nash

Nash isn't always pro-social!



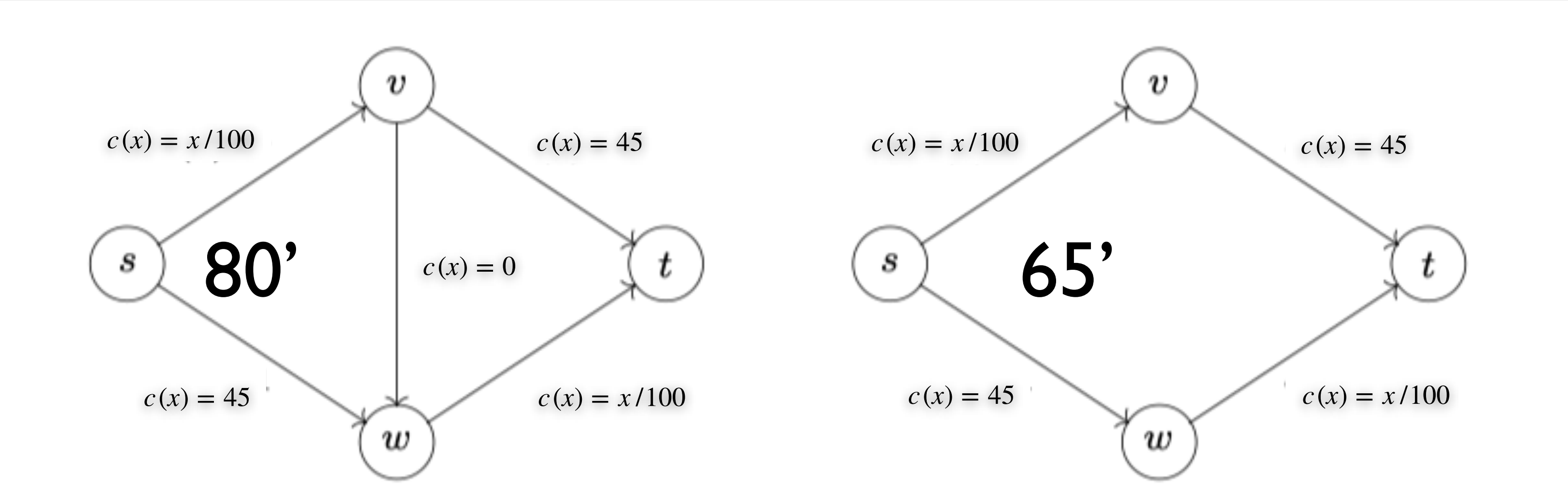
Removing a zero-cost edge improves social welfare!
Consider what is the Nash equilibrium (same travel time) for 4000 players

$$\text{right}: n = 4000 = x_{svt} + x_{swt} \wedge 45 + \frac{x_{svt}}{100} = 45 + \frac{x_{swt}}{100} \Rightarrow x_{svt} = x_{swt} = 2000 \text{ is Nash}$$

$$\text{left}: x_{svwt} = 4000 \text{ is Nash (and costs every player 15 minutes of travel time)}$$

Braess' Paradox *or* The problem with Nash

Nash isn't always pro-social!



Removing a zero-cost edge improves social welfare!
Consider what is the Nash equilibrium (same travel time) for 4000 players

$$\text{right}: n = 4000 = x_{svt} + x_{swt} \wedge 45 + \frac{x_{svt}}{100} = 45 + \frac{x_{swt}}{100} \Rightarrow x_{svt} = x_{swt} = 2000 \text{ is Nash}$$

$$\text{left}: x_{svwt} = 4000 \text{ is Nash (and costs every player 15 minutes of travel time)}$$

Learning in Populations games

Myopic best response is one particular form of how a population can learn "to play together."

What happens if we apply our previous theory of learning (reinforcement learning) in a population of agents?

Cross Learning - Stateless RL

- simplest form of RL
- each player *i* keeps a probability for each of their possible actions to execute it.
- agents reinforce these probabilities in the simplest possible way:
  - If Player R picks action *j* in the *n*-th interaction encountering a player playing action *k*, the probability of *j* for *R* is reinforced proportionately to the utility (payoff) experienced;
  - Any action not executed, has its probability decreased accordingly.

$$P_j(n+1) = U^R_{jk} + (1 - U^R_{jk})\, P_j(n)$$

$$P_{j'}(n+1) = (1 - U^R_{jk})\, P_{j'}(n) \qquad \text{for all} \quad j' \neq j.$$

Börgers and Sarin, Learning Through Reinforcement and Replicator Dynamics, Journal of economic theory 77, 114 (1997)

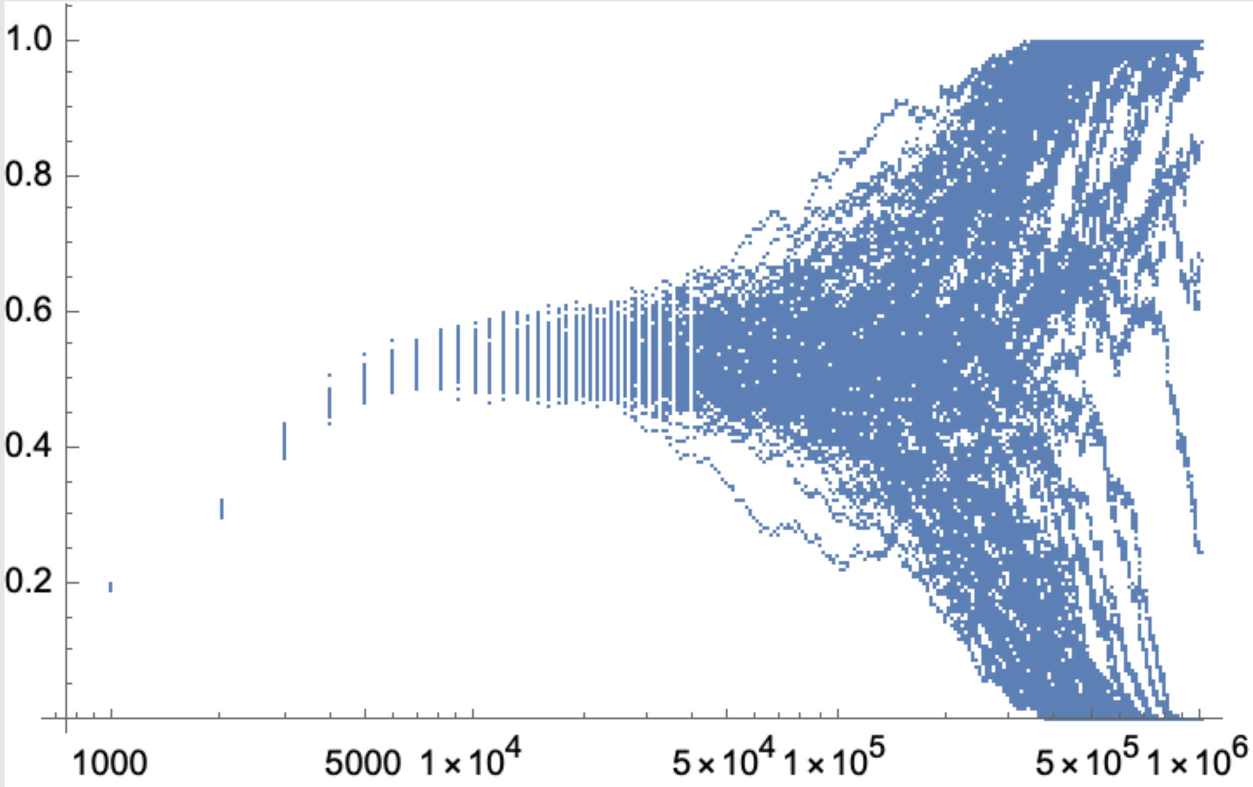
Cross Learning Dynamics

For a two-action system, the player only needs to keep track of one probability since $p_2^i = 1 - p_1^i$. Let the second player's action be *k*:

$$p_1^i \leftarrow U_{1,k} + (1 - U_{1,k}) \cdot p_1^i$$

i.e this probability is interpolated towards *p*=1 by a factor proportionate to the payoff.

Cross-Learning Example: Car Park Problem



scatterplot
probability
for each individual
(whole population)
against time

Students and tutors are always trying to come to arrive for class just in time. Unfortunately, this means that everyone arrives at the car park at the same time! There are two different car parks available. Car park A is large but far away from the building. Car park B is small but close to the building. It takes only 2 minutes to walk to the tutorial rooms from car park B but 20 minutes from A. Unfortunately, it is never easy to find a parking spot straight away and the situation gets worse the more people are trying to find a spot. If n cars arrive simultaneously at the large car park A, each will suffer an n-minute delay. In the smaller car park B, each car will even suffer a delay of $5n$-minutes if n cars are simultaneously looking for a spot.

Interpretations of Mixed Strategies

- Randomize individually

  Frequencies in repeated games individually

- Frequencies of occurrence in a population

  ie probability to meet certain opponents in a population game.

Social Learning in Populations games

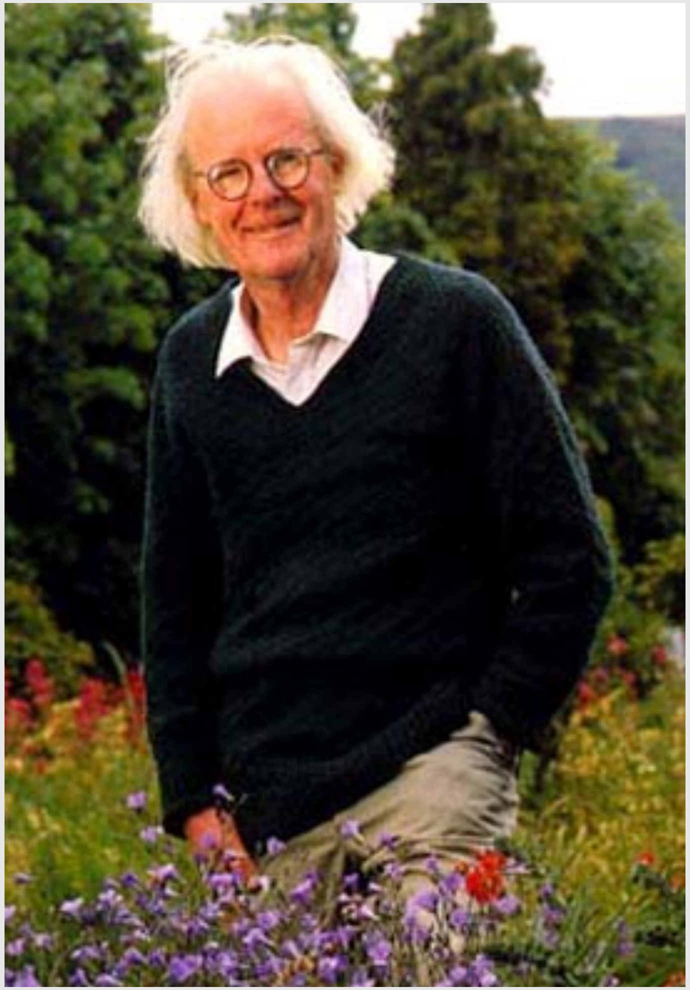
How can we understand the dynamics of this?

Much of our knowledge about learning in populations originates from evolutionary biology.

Evolutionary biology asks how species and populations adapt over time to the environment.

While this is a (subtly) different setting from learning much of this theory can be applied to learning in populations.

The Father of Evolutionary Game Theory



Maynard Smith (1986)
Evolution and the Theory of Games

Applied game theory to biological systems

Blind process of natural selection removes rationality assumption

"*Let nature do the work and see what happens*"

Results as if agents were rational, self-interested fitness maximizers

(but this is *not* by intentional design!)

EGT Departures from Classical Game Theory

- Population-level thinking (rather than two players)
- Interpretations as evolutionary or social process
- Games are repeated
- Stochastic model

- can also be interpreted as LEARNING !!
- Learning / imitation of behaviour
- Rationality not presumed
  - replaced by revision protocol -> later
- Perfect Information (underlies rationality) not required

Domains

- Biology
  - evolution (this is the origin)
  - group behaviour
- Economics
  - behavioural economics
  - e.g. public goods games
- Social Behaviour
- Psychology
- History (Cultural Evolution)
- …

Key Questions

- Which strategies “survive”, which vanish?
  - as agents learn from each other (social learning)

- What is the strategy distribution
  For a population in a stationary state (equilibrium)?

  - convergence on a single strategy or
    coexistence of different strategies?
  - may not be convergent but enter limit cycle
  - can a new strategy variation invade
    - (mutation/innovation)
    - can it displace the old strategy(ies)?

Evolution,

Learning …

General EGT framework

1. **Population of players $P_1, \ldots, P_m$**
2. **Pairs of players P_i, P_j meet and face-off ("play the game")**
   - **extensible to *n*-player games**
3. **Players receive payoff**
   - **viewed as fitness in the evolutionary setting**
4. **Strategies spread according to various mechanisms,**
   1. **e.g their relative fitness/performance (*details later*)**
5. **Game is repeated**

➡ **Extensive analytic theory for simple cases, in complex cases usually individual-based numerical models required**

First Analysis: *Assumptions*

- **Games are repeated.**
- **Strategies spread according to (proportional to) success.**
  - **This could be evolution (survival-of-the-fittest) or <u>social learning</u>**

- **Players are interchangeable and indistinguishable, except for strategy (as in population games)**

- **this implies that all players have the <u>same strategy set</u>**
- **it also implies that the game must be <u>symmetric</u>**
  - **i.e. the game is agnostic about the players’ roles**

Symmetric Games

Players are interchangeable, i.e. $a_{i,j} = b_{j,i}$ or

$$A = B^T$$

Since players are interchangeable we only consider symmetric pairs (x, x) for an equilibrium.

A strategy x is a Nash equilibrium if for all other strategies z:

$$x^T A x \geq z^T A x$$

It turns out that every symmetric game has a symmetric Nash Eq

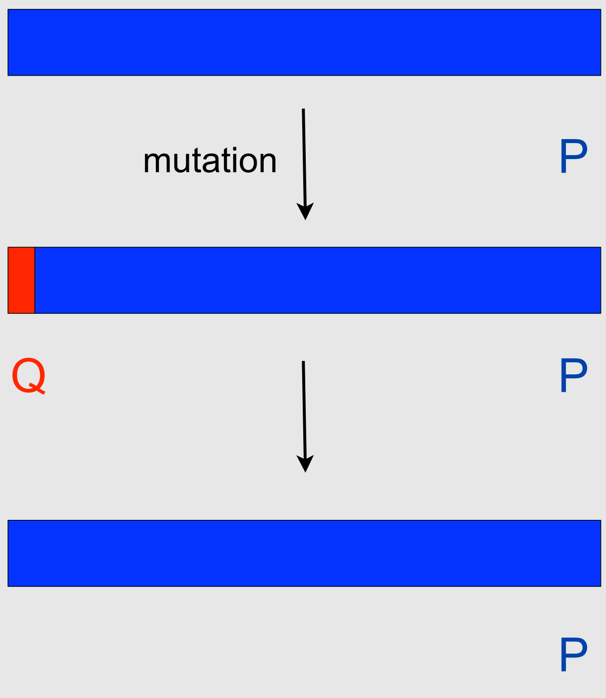
Nash Equilibrium & Process Dynamics

A dynamic process that always moves in the direction of increased payoff can only stop at a Nash Equilibrium

Can a mutant overturn a Nash Equilibrium?

Evolutionary Stable Strategy

- Imagine a <u>large population</u> of players, randomly drawn to play a symmetric game
- The payoffs of the game are assumed to be "fitnesses" in the sense that a process of (natural) selection favours those earning higher payoffs
- What is the condition for selection to <u>oppose the invasion of a mutant</u>?



P is ESS if it is able to repel any mutant Q that comes in a small quantity

Competition for food

- **When beetles of the same size compete, they get equal shares of the food**
- **When a large beetle competes with a small beetle, the large beetle gets the majority of the food.**
- **In all cases, large beetles experience less of a fitness benefit from a given quantity of food, since some of it is diverted into maintaining their expensive metabolism**

Body Size Game

		Beetle 2	
		Small	Large
Beetle 1	Small	5, 5	1, 8
	Large	8, 1	3, 3

Body Size Game: Invasion of a small mutant

- **Fraction of small Beetles: x**
- **Fraction of large Beetles: 1-x**

- **payoff in a population where Large is the majority (x is close to 0)**

Large Beetle: 8 * x + 3 * (1-x) = 3 + 5x ~ 3
Small Beetle: 5 * x + 1 * (1-x) = 1 + 4x ~ 1

- **Large is evolutionary stable**

Body Size Game: Invasion of a large mutant

- **Fraction of large Beetles: x** ***(swapped)***
- **Fraction of small Beetles: 1-x**

- **payoff in a population where Small is the majority (x is close to 0)**

Small Beetle: 5 * (1-x) + 1 * x = 5 - 4x ~ 4
Large Beetle: 8 * (1-x) + 3 * x = 8 - 5x ~ 8

- **Small is *not* evolutionary stable**

MONASH University
Information Technology

Evolutionarily Stable Strategy

P is ESS if ...

Nash Equilibrium → $\prod(P,P) \geq \prod(Q,P) \quad \forall Q$

if $\prod(P,P) = \prod(Q,P) \rightarrow \prod(P,Q) > \prod(Q,Q) \quad \forall Q \neq P$

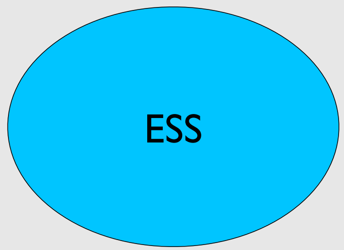
or equivalently,

Since expected payoffs can be linearly interpolated:

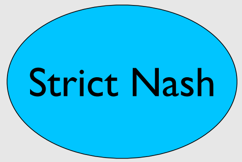
$$\prod(P, \epsilon Q + (1-\epsilon)P) > \prod(Q, \epsilon Q + (1-\epsilon)P) \quad \forall Q \neq P, \text{ small } \epsilon$$

"Even if the payoff is the same for P and Q in the all-P context, at some point P must be better off again as more Q move into the population"

ESS is a refinement of Nash



Strict Nash implies an ESS



ESS - Existence

Let G be a two-player symmetric game with 2 pure strategies such that

$$\Pi(s_1, s_1) \neq \Pi(s_2, s_1) \wedge \Pi(s_2, s_2) \neq \Pi(s_1, s_2)$$

then G has an ESS.

Replicator Dynamics

- **A dynamic systems description of evolutionary dynamics**
- **Replicator dynamics assumes exact copying of the behaviour proportional to fitness / success**
- **other types of learning / imitation (and thus dynamics) have been investigated**

Replicator Dynamics

large population, n types

x_i frequency of type i

$x = (x_1, \ldots, x_n) \in S_n$ state of population ←fraction of type *i*

$(Ax)_i$ average payoff for type i

$x^T Ax$ mean payoff in population

Important note: we are dealing with a discrete set of (immutable) strategy types here

Replicator equation

$$\dot{x}_i = x_i((Ax)_i - x^T Ax)$$

"x_i grows at the rate at which it outperforms the average"

$$\dot{x}_i / x_i$$ per capita rate of growth

adding constants to columns of A leaves equilibria unchanged!

Replicator dynamics - Rest Points

$$\dot{x}_i = x_i((Ax)_i - x^T Ax)$$

$$\text{rest point}: \ z_i = 0 \quad \text{or} \quad (Az)_i = z^T Az$$

Intuitively: either

- there is no-one left to execute strategy *i* or
- the payoff for strategy *i* is the same as the average payoff

This means that in steady-state all types must have same payoff.

Replicator equation for n=2

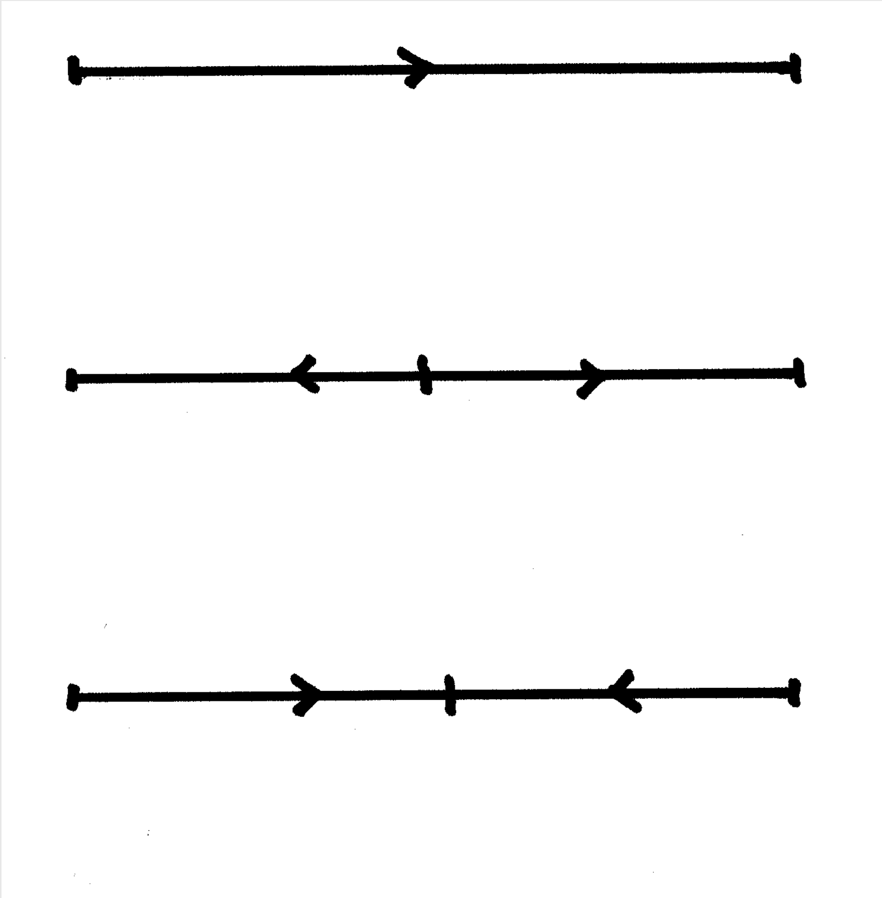
$$x_1 = x, \qquad x_2 = 1 - x$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{or equivalently} \quad \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$$

$$\dot{x} = x(1-x)[a-(a+b)x]$$

rest points for $x = 0, x = 1$ and $x = \dfrac{a}{a+b}$

Replicator equations for n=2:



- ✓ Dominance
- ✓ Bistability
- ✓ stable coexistence

Example coexistence: Hawk-Dove Game

$$\text{Hawk: escalate until injury}$$

$$\text{Dove: display, run away if co-player escalates}$$

$$G \text{ gain of contest}, -C \text{ fitness cost}, C > G$$

Type of adversary

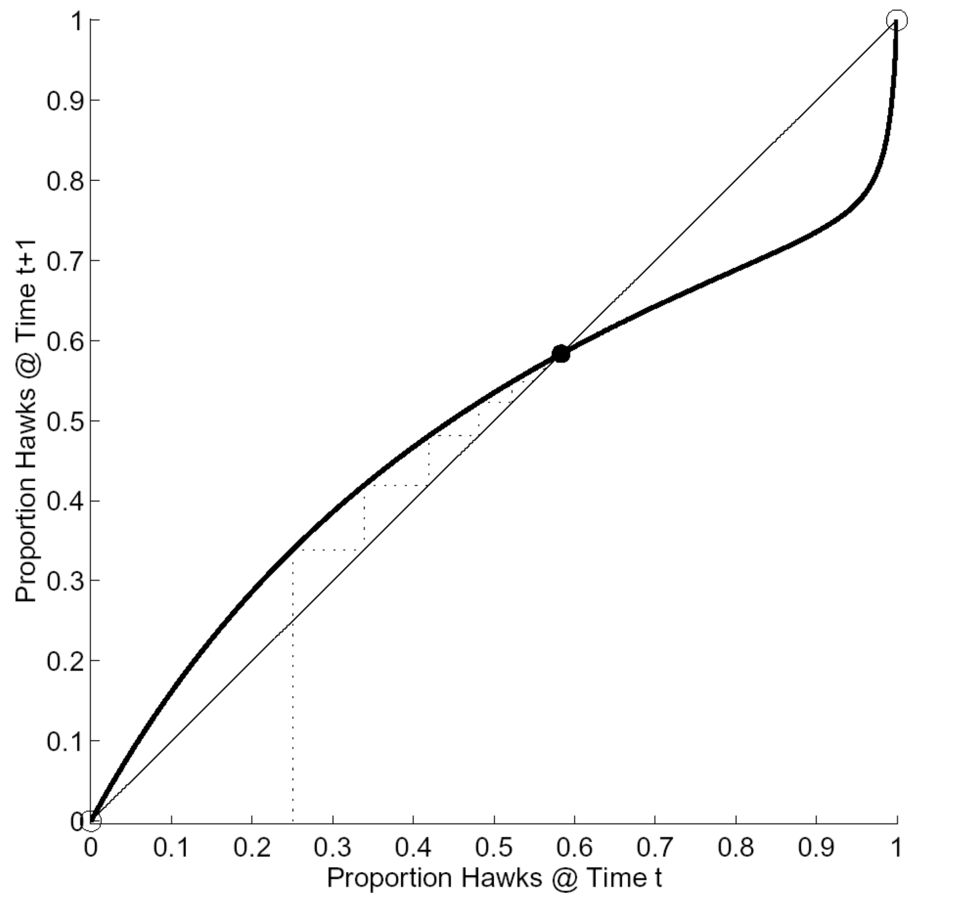
Contestant who wins payoff	Hawk	Dove
Hawk	(G-C)/2	G
Dove	0	G/2

Current Biology

$$x \to G / C$$

prevalence of ritual contests

Stable Polymorphic State



$$\begin{array}{c|c} \frac{G-C}{2} & G \\ \hline 0 & G/2 \end{array}$$

$$\Downarrow$$

$$\begin{array}{c|c} 0 & G/2 \nearrow a \\ \hline b \swarrow \frac{C-G}{2} & 0 \end{array}$$

$$\frac{a}{a+b} = \frac{G/2}{\frac{C-G}{2} + \frac{G}{2}} = G/C$$

Example dominance: Prisoner's Dilemma

$$\begin{bmatrix} 10 & -5 \\ 15 & 0 \end{bmatrix}$$

or equivalently

$$\begin{bmatrix} 0 & -5 \\ 5 & 0 \end{bmatrix}$$

$$x \rightarrow 0 \quad \text{(freq. of cooperators)}$$

Homework: check this by simulating RD

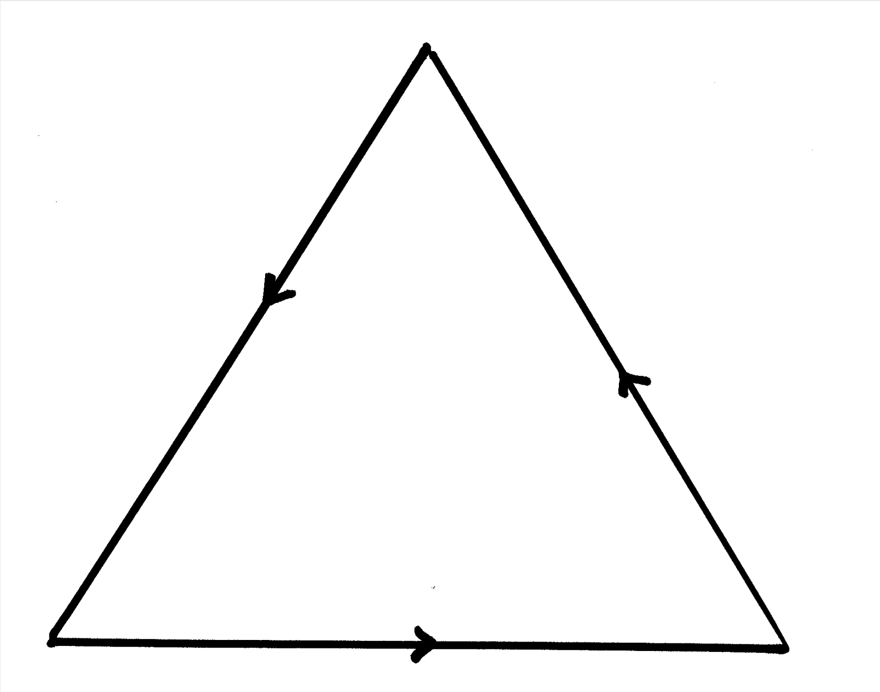
Link to Reinforcement Learning

Cross-Learning provably exhibits replicator dynamics. [in expectation and in the continuous time limit]

This is non-trivial: we now have a simple model of **individual learning** (based on reinforcement) that will follow the same dynamics as **social learning** (based on imitation).

This means, it will reach Nash under the same conditions etc.

Replicator equations for n=3:



Possibility of

heteroclinic cycles

(rock-scissors-paper)

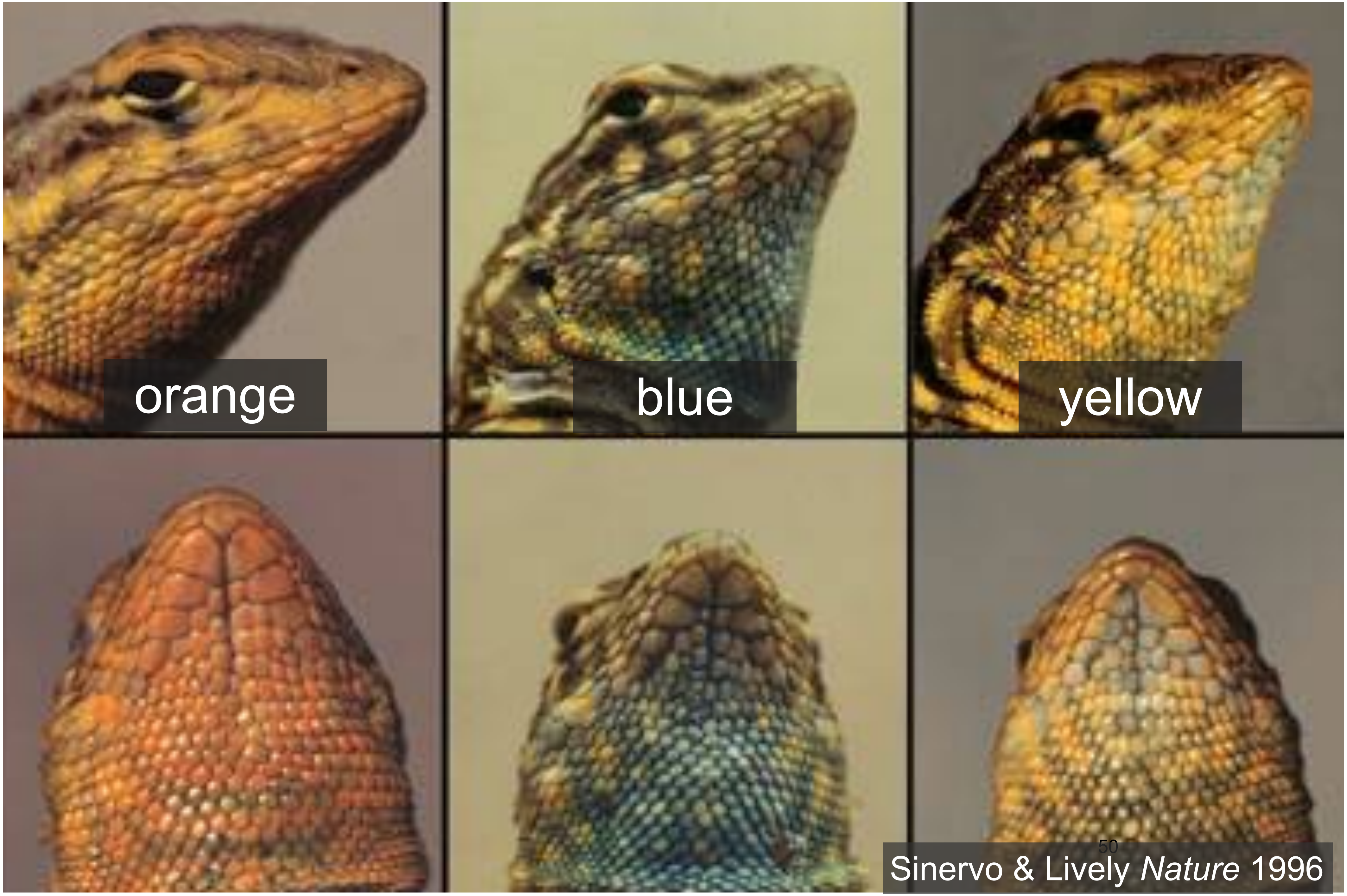
Rock-Scissors-Paper in nature

Uta stansburiana **(lizards)**

males: 3 morphs (inheritable)

1. **monogamous, guards female**
2. **polygamous, guards harem (inefficiently)**
3. **loose males, sneaky matings**

Male morphs in side-blotched lizard

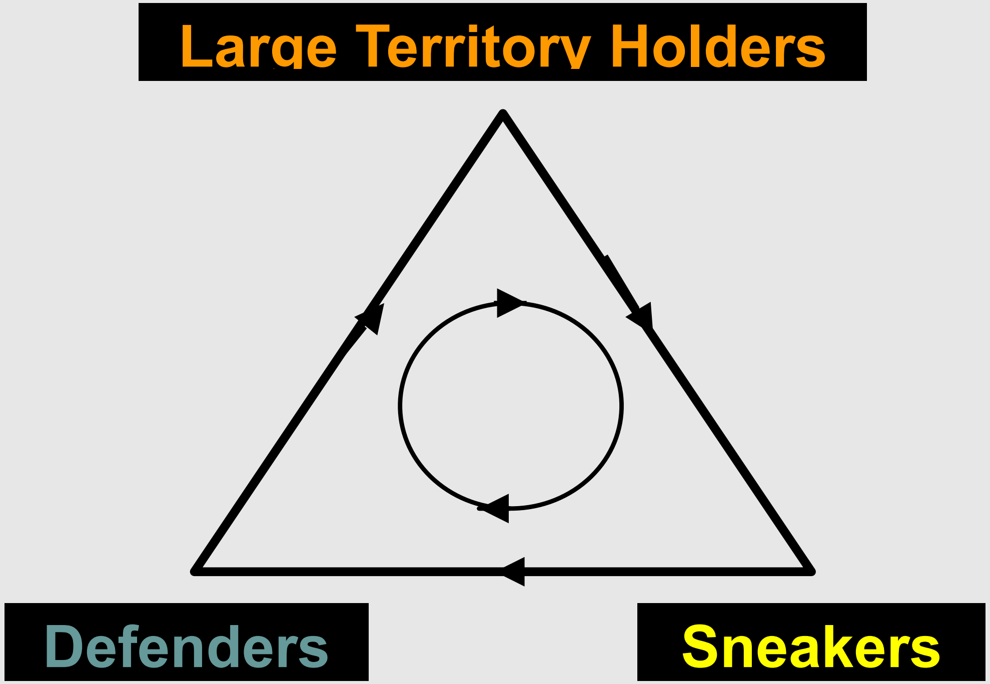


Sinervo & Lively *Nature* 1996

Evolutionary Game

Orange throated males: loosely defend large territory with several females
Yellow striped males: mimic females, can sneak into large territories
Blue throated males: firmly defend small territory with one female, not fooled by yellows

Evolutionary game:
- blue beats yellow,
- yellow beats orange,
- orange beats blue
- results in 5 year cycles



“Folk theorem” of evolutionary game theory

- ➡ **if the dynamics converges, it converges to NE**
- ➡ **strict NE are attractors**
- ➡ **stable rest points are NE**

The replicator equation and other game dynamics.
Ross Cressman and Yi Tao, PNAS 111:10810-10817

Continuous-valued games: Snowdrift

Freeloaders

Snowdrift Games

Similar to prisoner's dilemma
but here with continuous trait value x
(investment in the task)
and continuous payoff

$$\Pi(X, Y) = B(X + Y) - C(X)$$

collective benefit but individual cost

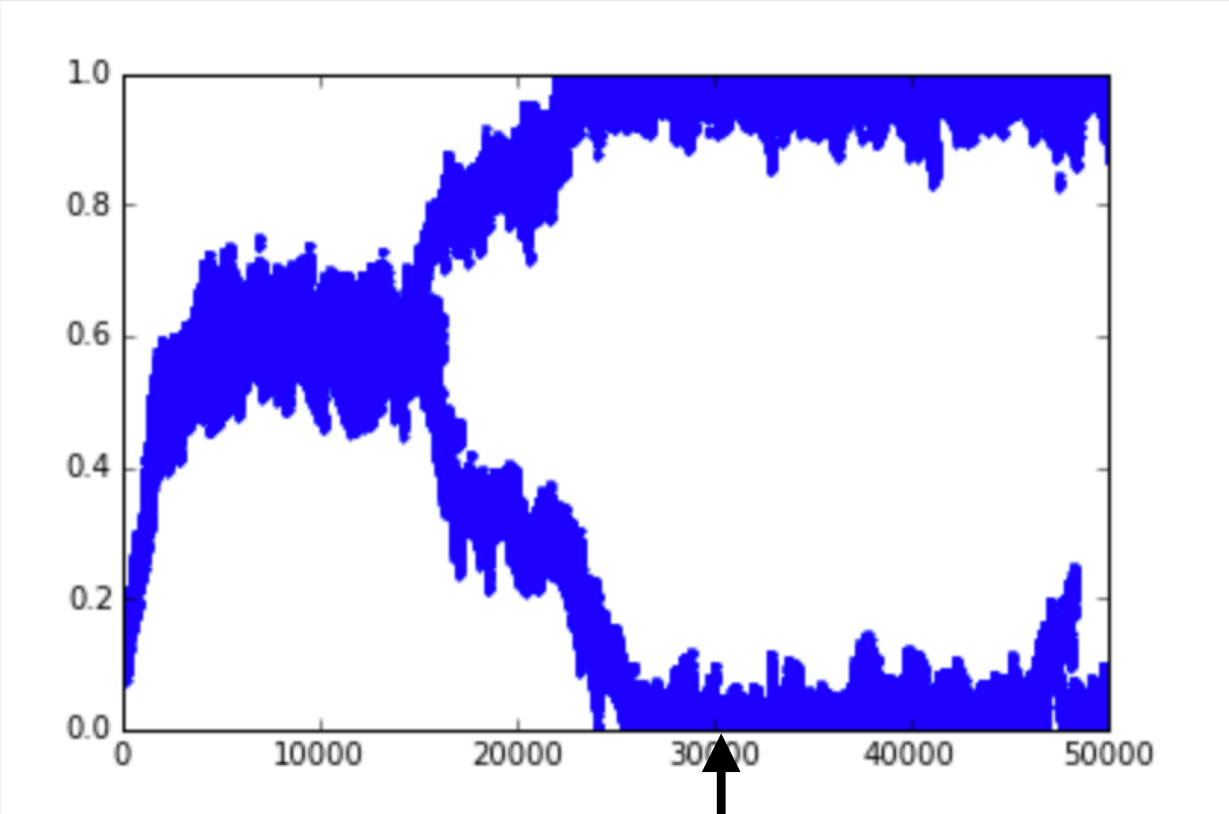
we can't handle continuous traits with replicator!

Replicator with Mutation: Snowdrift

```
def oneRound():
    global population
    opponents = np.array(population)
    np.random.shuffle(opponents)
    allGames=np.array([population, opponents]).T
    allPayoffs=map(payoffGame,allGames)
    parents=np.random.choice(population,popSize,
        replace=True,p=allPayoffs/np.sum(allPayoffs))
    population=np.array(map(mutateIf,parents))


def payoffGame(G):
    return benefit(G[0]+G[1])-cost(G[0])
```

Snowdrift Game Simulation



Freeloaders

Reminder: Interpretations of mixed Strategies

- Randomize individually
- Frequencies in repeated games individually
- Frequencies of occurrence in a population

  ie probability to meet certain opponents in a population game.

Adaptive dynamics framework

- can handle continuous traits (continuous strategy space) but:
- Population is assumed to be homogeneous (monomorphic): all players adopt same strategy

- **Mutation generates variant strategies very close to the resident strategy and is rare**
- **If a mutant beats the resident players it takes over otherwise it is rejected**

- **can explain branching**
  (but nothing after the branching point)

Some other Learning & Revision Protocols

- Pairwise Comparison / Smith Dynamics
  - *revising agents choose a candidate strategy at random, switching to it with positive probability if and only if its payoff is higher than the agent's current strategy.*
  - *Pairwise comparison induces Smith Dynamics, which is* Nash-stationary*: the set of rest points is identical to the Nash equilibria*

- Best Response Dynamics
  - *revising agents choose a set of opponents at random and switch to a best response strategy for these opponents.*

Take home lessons 1/2

- Evolutionary Game Theory is a useful tool for describing the learning of a group of agents. While it was originally designed to describe evolutionary processes it can be reinterpreted as describing learning by switching the timescales.

- It describes the development of strategies in a large population, typically in symmetric games and of interchangeable agents that all have the same action set.

- It addresses questions such as “what strategies can survive in a population?”; “can different strategies co-exist?” etc.

- EGT introduces a new outcome concept: the Evolutionary Stable Strategy (ESS). A strategy is evolutionarily stable if it cannot be displaced by small mutations in a population.

- $\text{strict NE} \subset \text{ESS} \subset NE$

- Every 2-player 2-strategy symmetric game has an ESS.

- Replicator dynamics describes how the proportions of different strategies in a population develop under such *social* learning (agents learning from each other). This can also be seen as evolutionary selection.

- Cross-learning is one of the simplest forms of reinforcement learning. It describes *individual* learning from experience. In the continuous-time limit and in expectation it converges to the same outcome as replicator dynamics.

Take home lessons 2/2

- Replicator dynamics can only describe agents with discrete action sets. Other (more complex) formalism for continuous action sets exist (e.g. adaptive dynamics).

- In replicator dynamics, agents modify their strategies according to a so-called *revision* protocol. The most common revision protocol is imitation dynamics, in which agents copy strategies from each others with a probability that depends on the success of these strategies.

- Since agents only copy strategies, replicator dynamics (without mutation) can never produce a *new* strategy. It only describes how proportions of existing strategies grow and shrink.

- Other protocols are possible (e.g. best response, pairwise comparison) and many result in very similar dynamics.

- Replicator dynamics can be described by the ODE system $\dot{x}_i = x_i((Ax)_i - x^T Ax)$

- It can also be analysed using agent-based simulation.

- Under imitation dynamics, replicator dynamics converges on Nash Equilibria ("Folk Theorem").

- In a 2-strategy game only three outcomes are possible: either one of the strategies dominates or there is a point of stable coexistence. If there are more than 2-strategies, the dynamics becomes far more complex (e.g. cycling is possible).