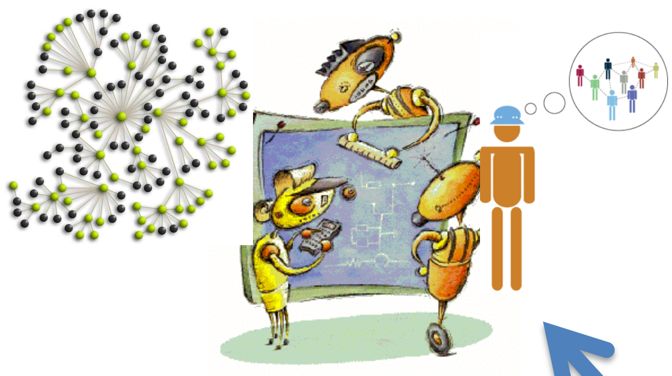
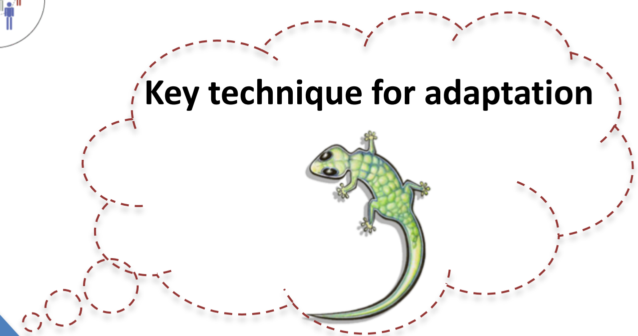
COMP 532
Machine Learning and BioInspired Optimization

Lecture 19: Multi-Agent Learning

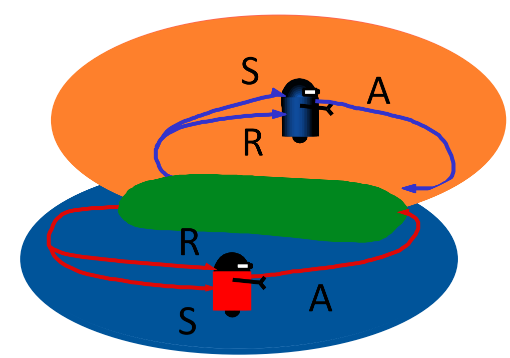
Dr. Shan Luo
Department of Computer Science
shan.luo@liverpool.ac.uk

Remember..?

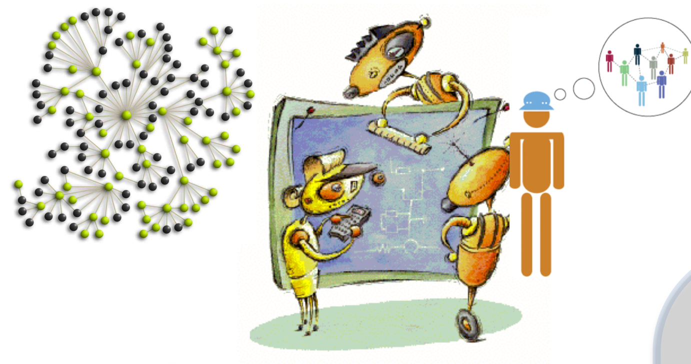




Reinforcement Learning



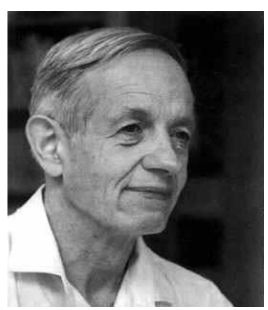
Remember..?



Strategic decision making

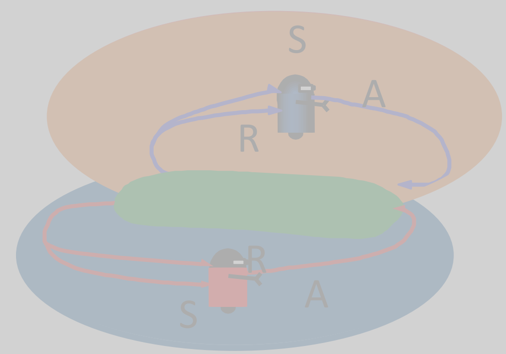
Classical Game Theory

- CENTRAL CONCEPT:
  Nash equilibrium
- Normative theory
- Rational players

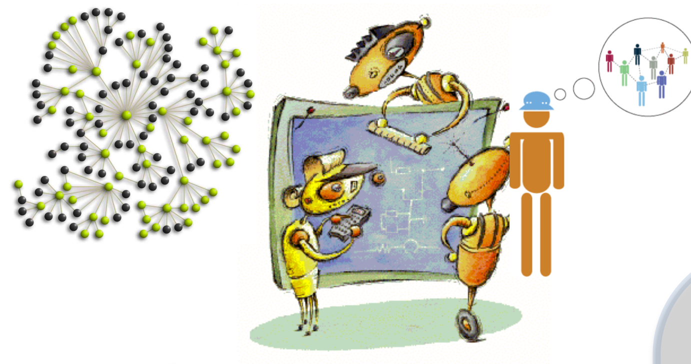


Key technique for adaptation

Reinforcement Learning



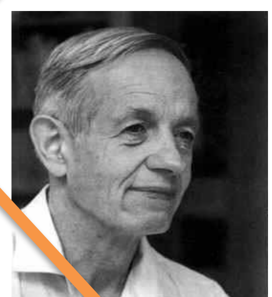
Remember..?



Strategic decision making

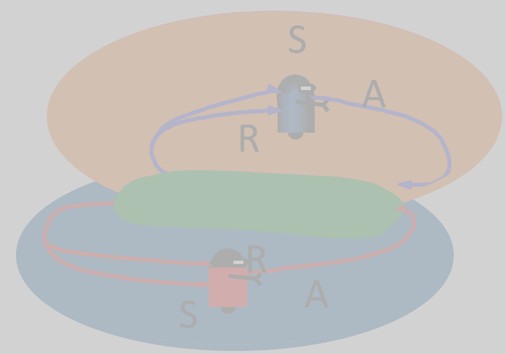
Classical Game Theory

- CENTRAL CONCEPT:
  Nash equilibrium
- Normative theory
- Rational players

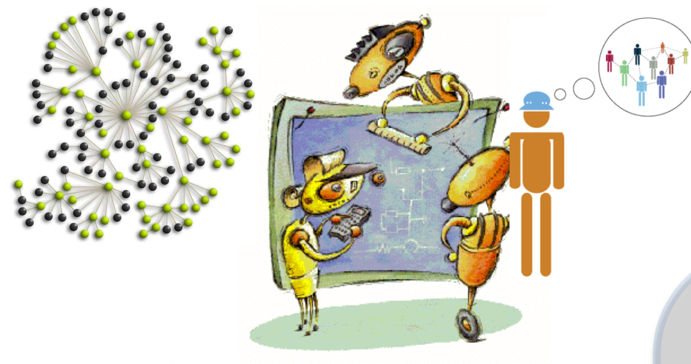


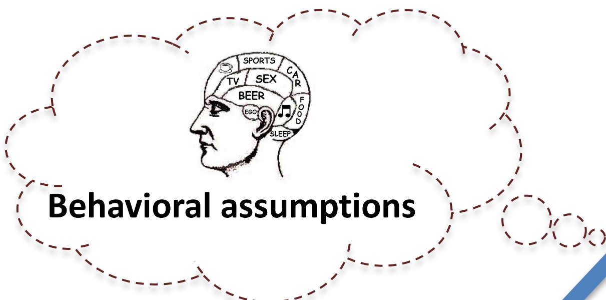
Key technique for adaptation

Reinforcement Learning

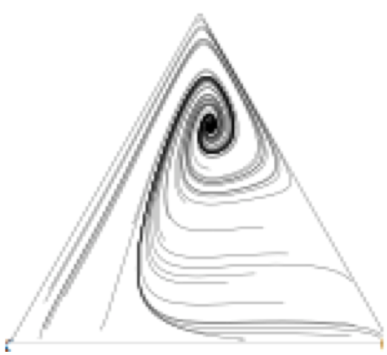


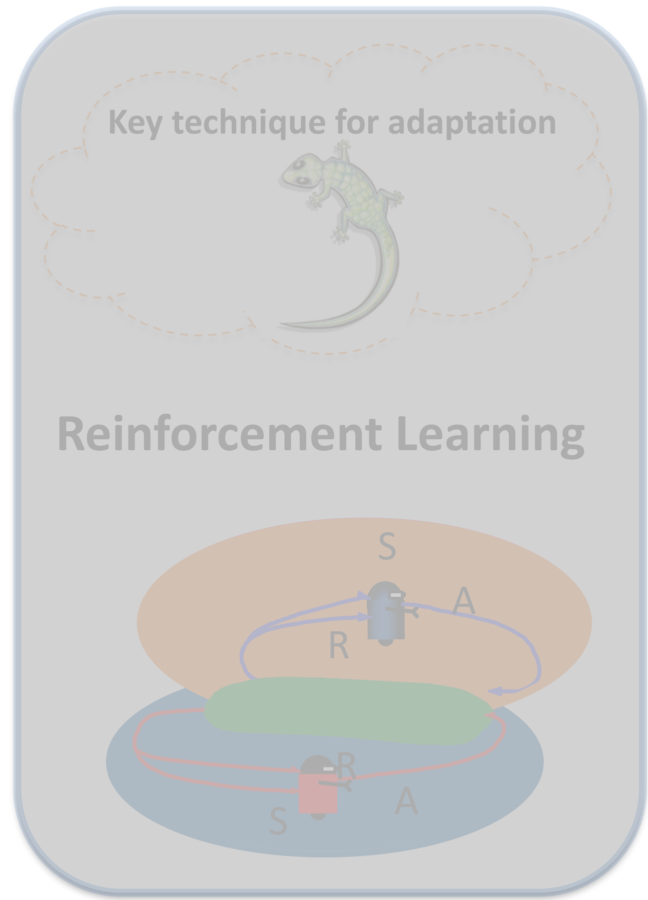
Remember..?



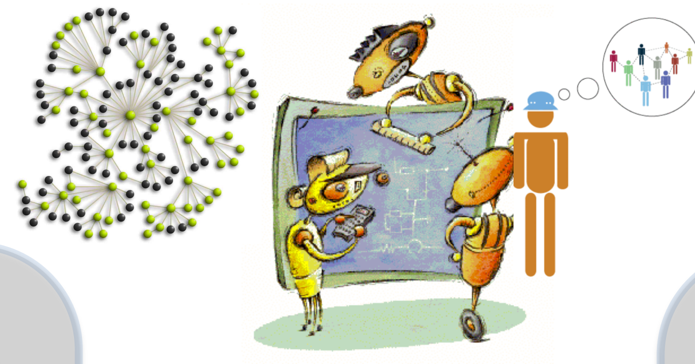


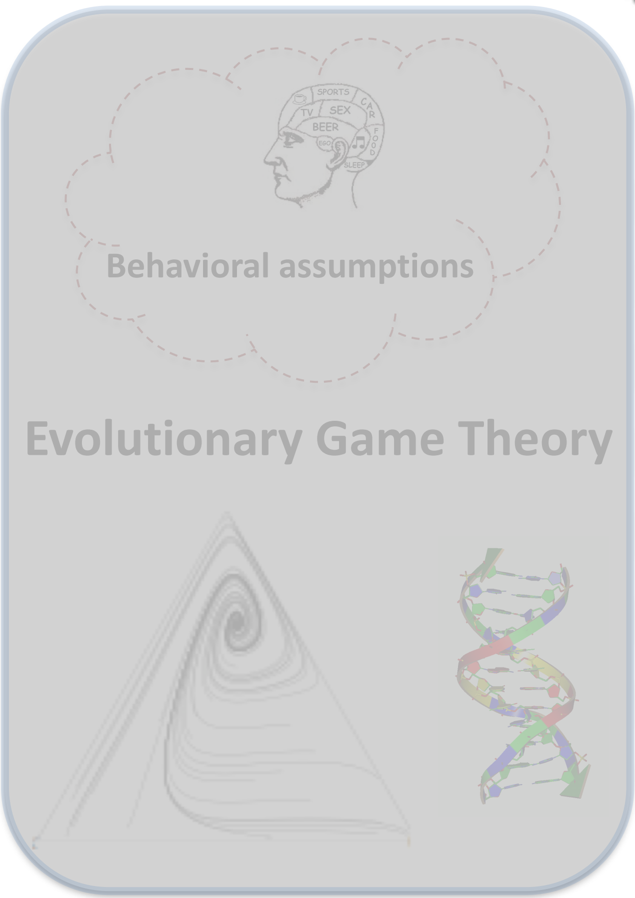
Evolutionary Game Theory





Remember..?

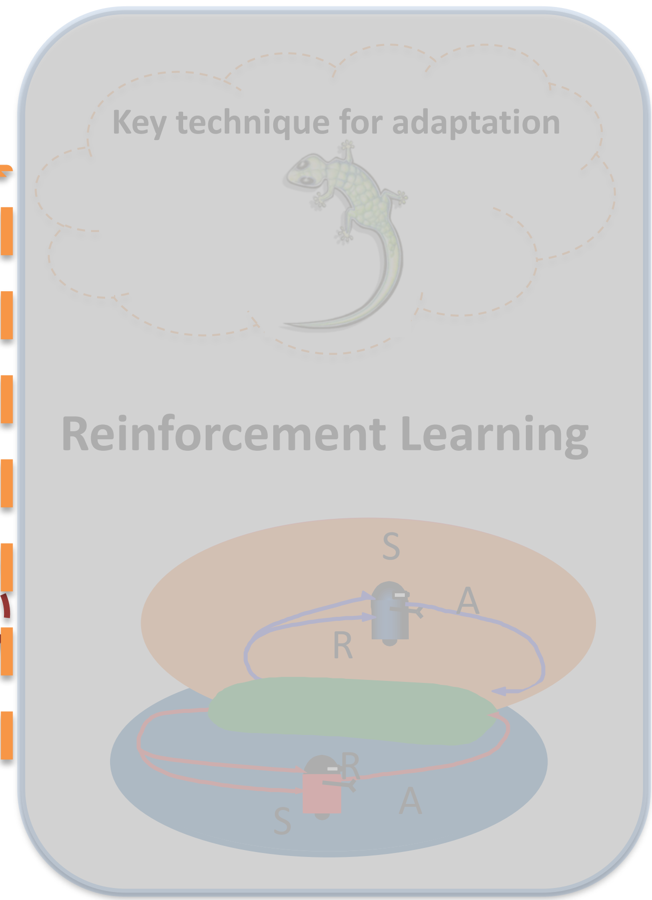




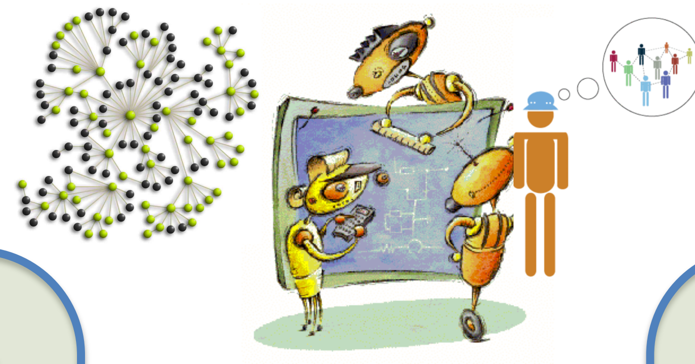
Backbone

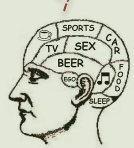
$$\frac{dp_i}{dt} = p_i\left[e_i A q - p A q\right]$$

$$\frac{dq_i}{dt} = q_i\left[p B e_i - q B p\right]$$



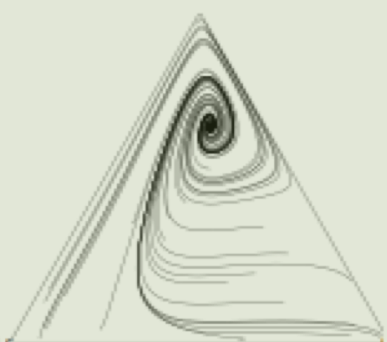
Remember..?

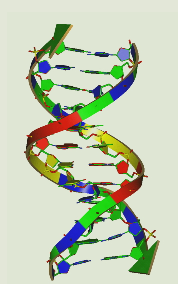




Behavioral assumptions

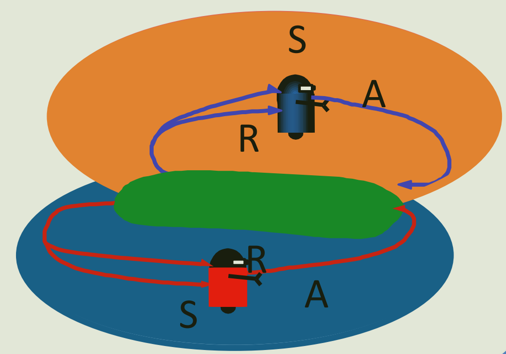
Evolutionary Game Theory





Key technique for adaptation

Reinforcement Learning

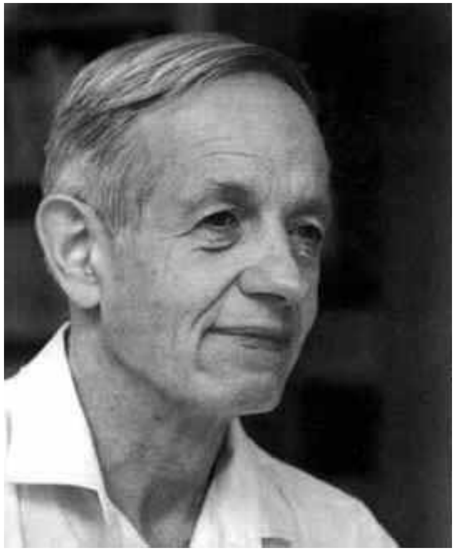


Outline (3-ish lectures)

- Introduction to Evolutionary Game Theory
  - Replicator Dynamics
  - Evolutionarily Stable Strategies
  - Example games
- Formal link between RL and EGT
  - Deriving the dynamics of Cross Learning
  - Extension to other RL algorithms
- Applications of this model
  - Parameter tuning
  - Analyzing complex strategic interactions

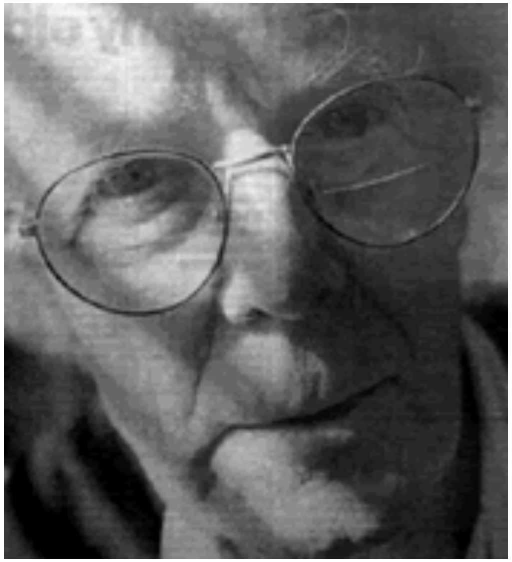
Back to Game Theory

- Classic:
  - Economical theory (von Neumann, Morgenstern, later Nash)
  - Normative Theory
  - Modelling interactions through games
  - Central concept: **Nash equilibrium**



Back to Game Theory

- Evolutionary (John Maynard-Smith):
  - Based on biological evolution
  - Descriptive theory
  - Strategies evolve through selection and mutation
  - Central concepts:
    Evolutionarily Stable Strategies
    Replicator Dynamics



Evolutionary Game Theory

- Studies a **population**..

  ..of individuals of different **types**..

  ..who are randomly paired in interaction..

  ..and whose relative **fitness** determines their reproductive success

- Evolutionary operators:
  - **Selection** ↔ exploitation
  - **Mutation** ↔ exploration

Example: Prisoner's Dilemma

Remember..

- Two players:
  - Row player
  - Column player
- Actions:
  - Cooperate = hold out
  - Defect = confess
- Nash equilibrium:
  - Mutual defection

	C	D
C	3, 3	0, 5
D	5, 0	1, 1

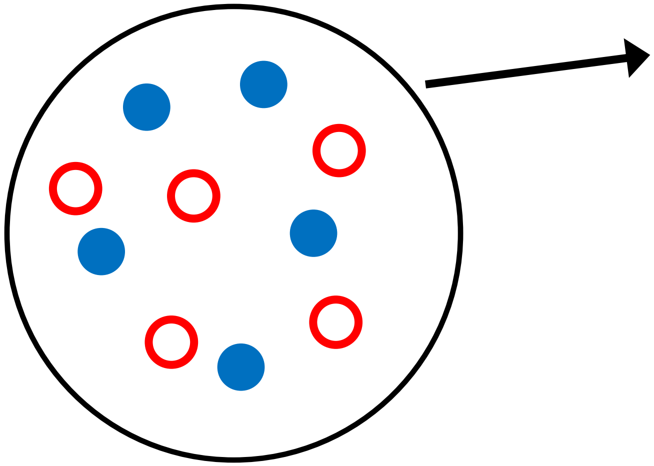
Example: Prisoner's Dilemma

Evolutionary Game Theory..

- A population of individuals
- Types:
  - Cooperators
  - Defectors
- Individuals are paired randomly
  - Combination of types determines their reward = fitness

	C	D
C	3	0
D	5	1

Example: Prisoner's Dilemma



Population

● = Cooperator

○ = Defector

Assumption:
Infinite population!

Interactions:

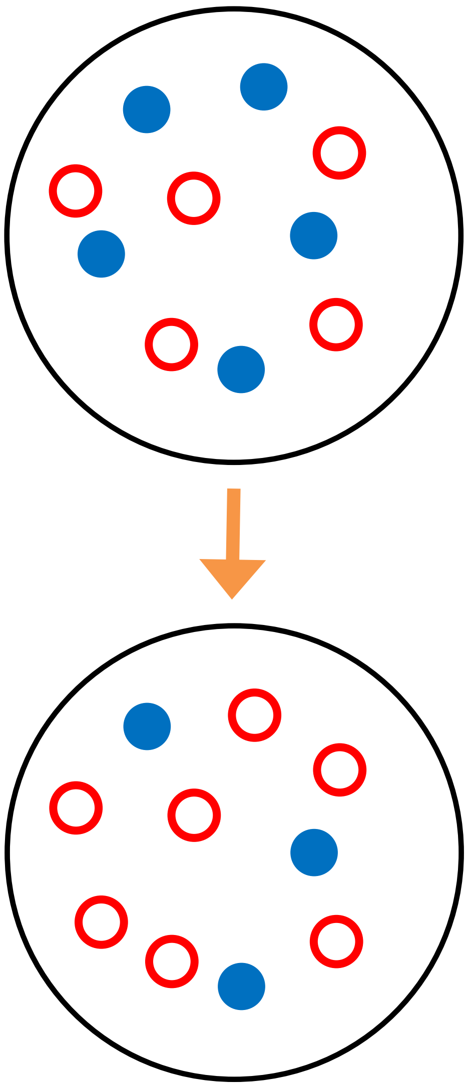
● + ● → both get a fitness of 3

○ + ○ → both get a fitness of 1

● + ○ → ● gets 0 and ○ gets 5

	C	D
C	3	0
D	5	1

Example: Prisoner's Dilemma



- Suppose #cooperators = #defectors
- All pairings have equal chance
  - Defectors on average get a fitness of $(5{+}1)/2 = 3$
  - Cooperators get $(3{+}0)/2 = 1.5$

→ **Defectors have a higher reproductive success**

	C	D
C	3	0
D	5	1

Replicator Dynamics

- The fitness of each type defines their reproductive success, and hence the proportional growth of that type in the population

- **Replicator Dynamics**:

$$\dot{x}_i = x_i \left[f_i(x) - \underbrace{\sum_j x_j f_j(x)}_{\text{average fitness of the population}} \right]$$

growth rate

fraction of type i in the population

fitness of type i

average fitness of the population

Replicator Dynamics

- **Replicator Dynamics**:

$$\dot{x}_i = x_i \left[f_i(x) - \sum_j x_j f_j(x) \right]$$

- Growth rate of type i is proportional to the difference between its fitness and the average fitness of the population
- **Intuition**: types that do better than average increase; types that do worse decrease

Back to the Prisoner's Dilemma

$$\dot{x}_i = x_i \left[f_i(x) - \sum_j x_j f_j(x) \right]$$

- In a population with the distribution over Cooperators and Defectors given by x, and payoff matrix A

Assume x is a column vector

$$f_i(x) = (Ax)_i$$

$$\sum_j x_j f_j(x) = x^{\top} A x$$

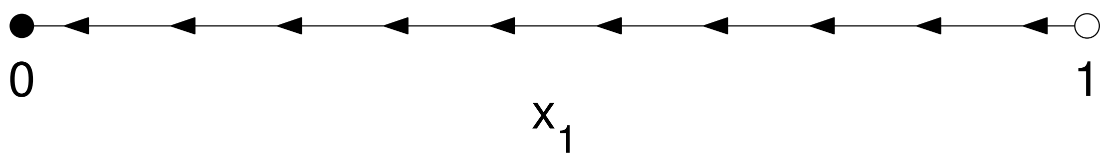
	C	D
C	3	0
D	5	1

$$\dot{x}_i = x_i[(Ax)_i - x^{\top} A x]$$

Back to the Prisoner's Dilemma

$$\dot{x}_i = x_i[(Ax)_i - x^{\intercal}Ax]$$

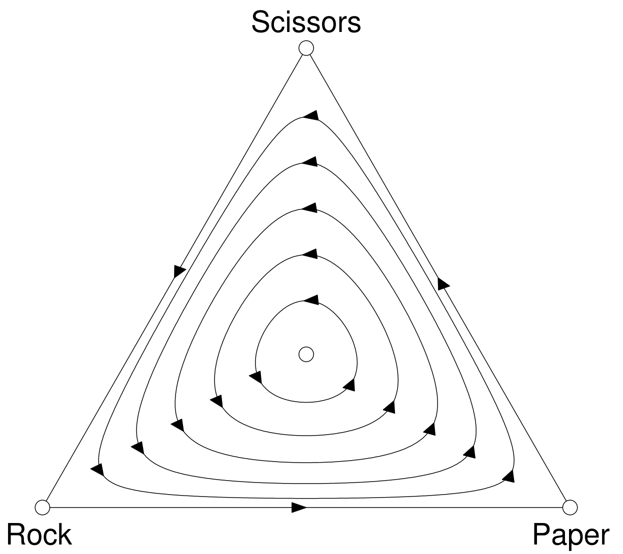
	C	D
C	3	0
D	5	1



All Defectors **All Cooperators**

Back to Rock-Paper-Scissors

	R	P	S
R	0	-1	+1
P	+1	0	-1
S	-1	+1	0



Single Population Replicator Dynamics

$$\dot{x}_i = x_i[(Ax)_i - x^\top Ax]$$

- So far we have modelled a single population
  - **Intuition**: a single player, optimising her strategy by playing against herself
- Typically we are interested in games in which multiple independent players interact
  - A different population for each player?

Multi-Population Replicator Dynamics!

Multi Population Replicator Dynamics

- **Two populations**
  - types distributed according to x and y
  - payoff matrices A and B
  - randomly pair one individual from each population

$$\dot{x}_i = x_i[(Ay)_i - x^\intercal Ay]$$

$$\dot{y}_i = y_i[(x^\intercal B)_i - x^\intercal By]$$

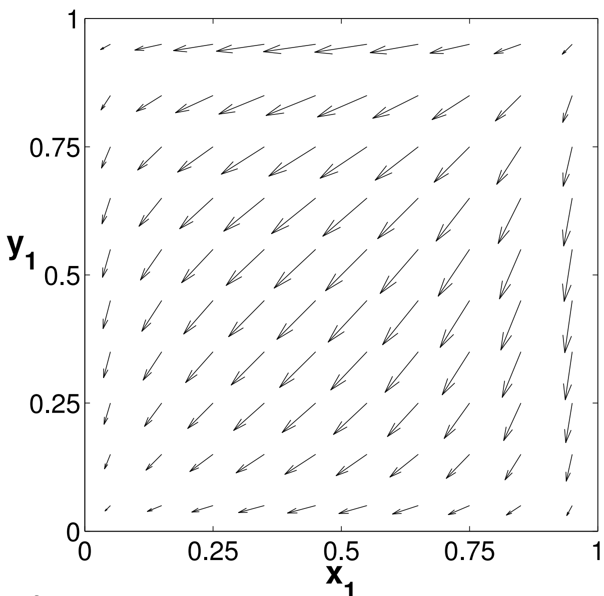
Back to the Prisoner's Dilemma

$$\dot{x}_i = x_i[(Ay)_i - x^\intercal Ay]$$

$$\dot{y}_i = y_i[(x^\intercal B)_i - x^\intercal By]$$

	C	D	
C	3, 3	0, 5	x_1
D	5, 0	1, 1	x_2
	y_1	y_2	

(D, C) (C, C)



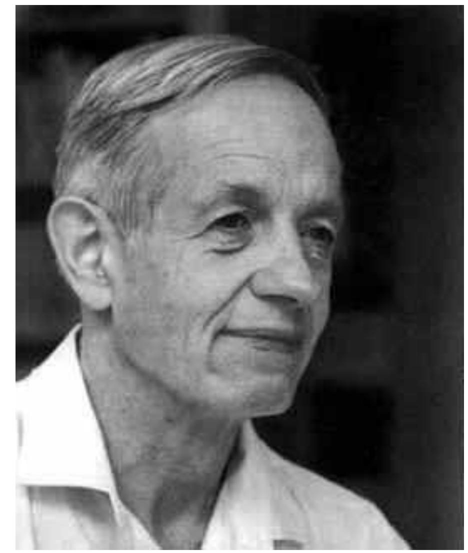
(D, D) (C, D)

Dictionary

Reinforcement Learning	Classical Game Theory	Evolutionary Game Theory
environment	game	game
agent	player	population
action	action	type
policy	strategy	distribution over types
reward	payoff	fitness

Remember: Nash equilibrium

- Best-response equilibrium
- Assumes rational players and full knowledge of the game
- Static concept: the final outcome

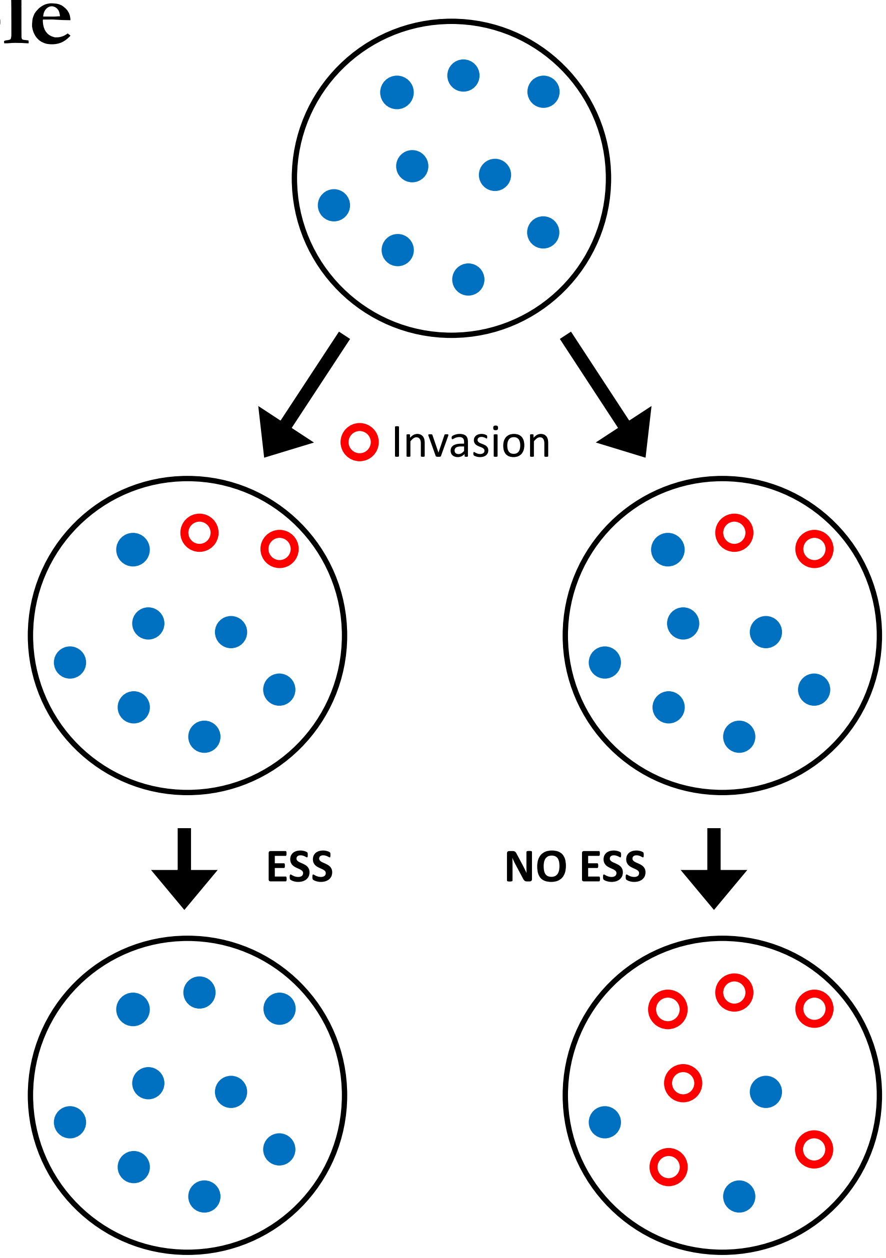


Intuitively: A Nash equilibrium is a strategy profile for a game, such that no player can increase her payoff by changing her strategy, while the other players keep their strategy fixed.

Evolutionarily Stable Strategies

Evolutionary Stable Strategies (ESS)

- John-Maynard Smith and Price (1973)
- Populations of individuals play particular strategy
- If an invasion of another strategy is impossible, then the strategy is ESS.



Evolutionarily Stable Strategies

- ESS are
  - Population distributions that are asymptotically stable fixed points of the replicator dynamics
  - Refinement of Nash equilibrium: $ESS \subseteq NE$

- Formally, strategy x is an ESS iff for any mutant y:
  1. $f(x,x) \geq f(y,x)$ and
  2. if $f(x,x) = f(y,x)$ then $f(x,y) > f(y,y)$

Back to the Prisoner's Dilemma

- Formally, strategy x is an ESS iff for any mutant y:
  1. $f(x,x) \geq f(y,x)$ and
  2. if $f(x,x) = f(y,x)$ then $f(x,y) > f(y,y)$

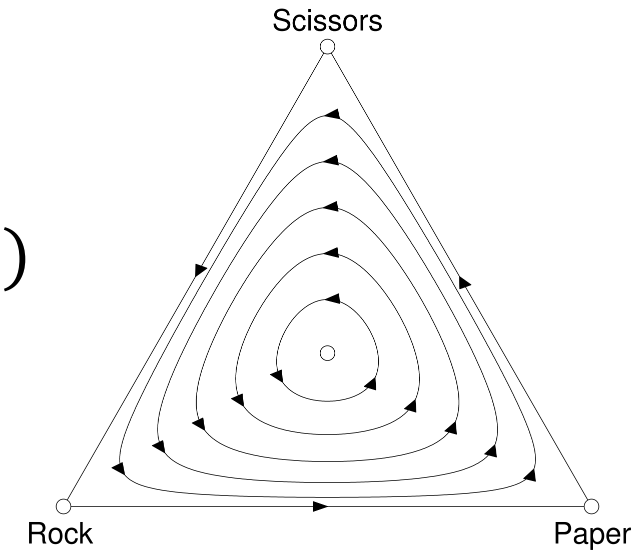
- In the Prisoner's Dilemma, we have
  - $f(D,D) > f(C,D)$ → done!
  - Mutual defection is ESS

	C	D
C	3	0
D	5	1

Back to Rock-Paper-Scissors

- ESS are
  - Population distributions that are **asymptotically stable fixed points** of the replicator dynamics

- Without doing the maths..
  - NE: $\Pr(R, P, S) = (1/3, 1/3, 1/3)$
  - Is this also an ESS?



Selection Mutation Dynamics

- So far we have only discussed **selection** dynamics
- We can also add **mutation**
  - E.g., suppose a mutation rate ε

$$\dot{x}_i = x_i\left[f_i(x) - \sum_j x_j f_j(x)\right] + \sum_j [\underbrace{\varepsilon x_j}_{\text{j mutates into i}} - \underbrace{\varepsilon x_i}_{\text{i mutates into j}}]$$

Next up…

- We will see that the replicator dynamics are not just useful for game theory

- They can also **predict learning behaviour** of reinforcement learning agents!

- We will formally derive this link, and look at a range of applications