Evolutionary games: competition and cooperation

Kristian Lindgren
Department of Energy and Environment
Chalmers, Göteborg, Sweden

- Model for evolution of behaviour in repeated Prisoner's Dilemma
  - Representation of agents: strategies, genes, ...
  - Discussion on simulation example
- On a fundamental problem with the Nash equilibrium concept
  - The backward induction paradox
  - How to model "rational" behaviour in agents?

Prisoner's Dilemma

Payoff matrix:

Action for player A		Action for player B: Cooperate	Action for player B: Defect
	Cooperate	(3,3)	(0,5)
	Defect	(5,0)	(1,1)

- Two actions: Cooperate and Defect
- Only one Nash equilibrium (situation where no player wants to change action): both players Defect

The repeated Prisoner's Dilemma

In the repeated game the same pair of players meet in a series of rounds.

Payoff matrix:

		Action for player B	
		Cooperate	Defect
Action for player A	Cooperate	(3,3)	(0,5)
	Defect	(5,0)	(1,1)

- A key strategy is **Tit-for-tat**:
  start with cooperation (C) then mimic opponent's previous move
- We assume a probability r for the game to end after each round: average game length $n = 1/r$.
- Tit-for-tat is a Nash equilibrium if the expected number of rounds n is large enough.

- We introduce mistakes as a complication. Can cooperation be established without being exploited? (mistake rate 1%)

Evolutionary model

- Fixed number of players (1000)
- Each player plays against everyone else, and each player accumulate their scores to form their *fitness*
- A genetic code determines the strategy for the player
- Replicator dynamics:
  - In each generation, a fraction δ of the population is replaced
  - Mutations
  - Fitness-proportionate selection (determined by score f_i for strategy of type i)

$$x_i' = (1-\delta)x_i + \delta \frac{f_i}{\bar{f}} x_i$$

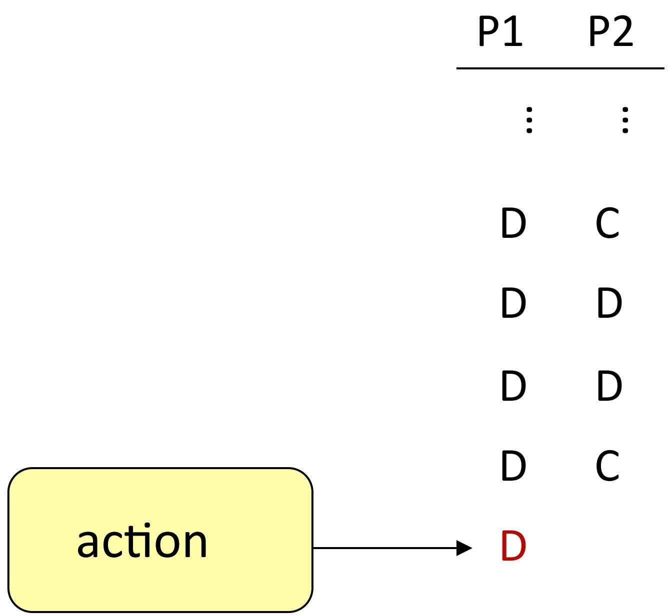
Models of behaviour in repeated games

Two players:	P1	P2
	⋮	⋮
	D	C
Game history:	D	D
	D	D
	D	C
Next action:	?	

C = Cooperate
D = Defect

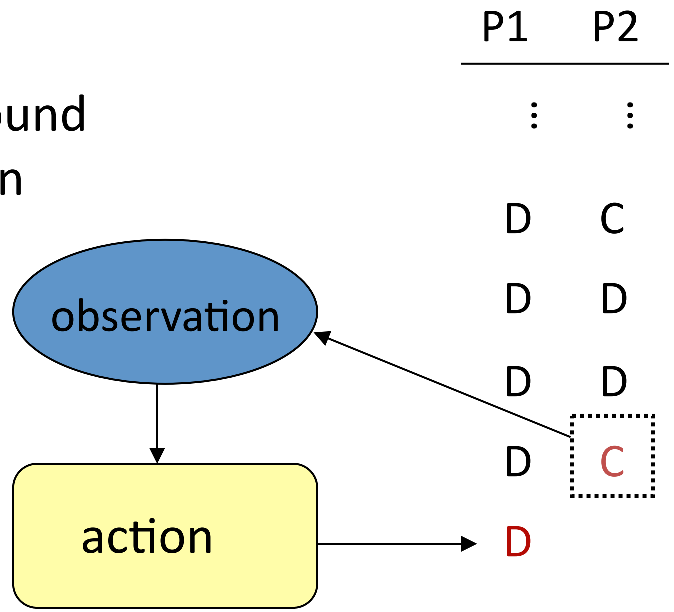
Models of behaviour in repeated games

- Strategy without memory



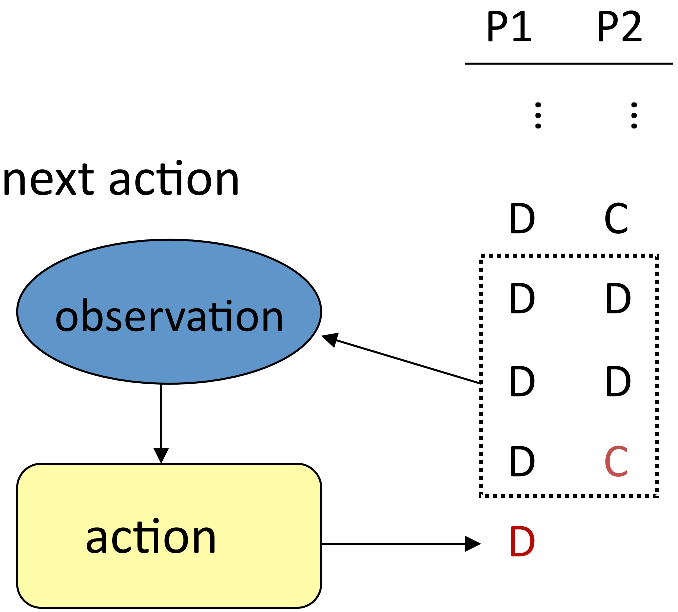
Models of behaviour in repeated games

- Reactive strategy
  opponent's action in last round
  determines your next action



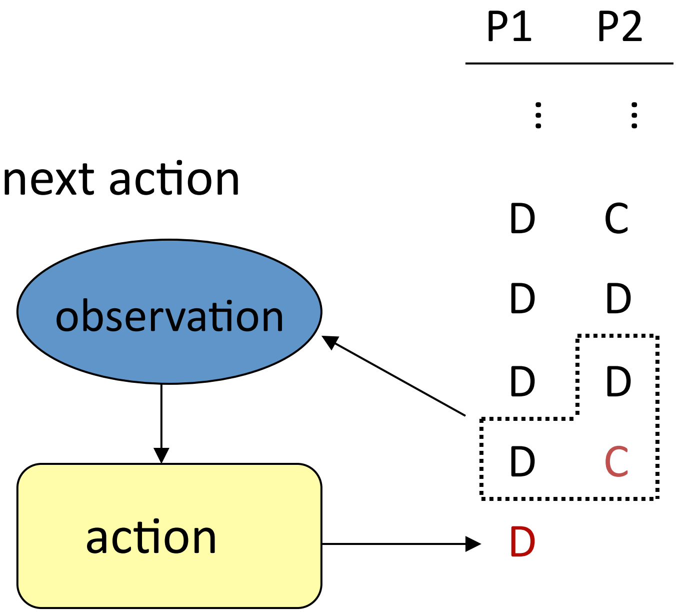
Models of behaviour in repeated games

- Reactive strategy
  more general:
  mapping a finite history to next action



Models of behaviour in repeated games

- Reactive strategy
  more general:
  mapping a finite history to next action



Example:
Look-up table & genetic coding

observation	action
D D D	D
D D C	D
D C D	D
D C C	C
C D D	D
C D C	C
C C D	D
C C C	C

The strategy is determined by the output column that consequently can serve as the genetic code: D D D C D C D C

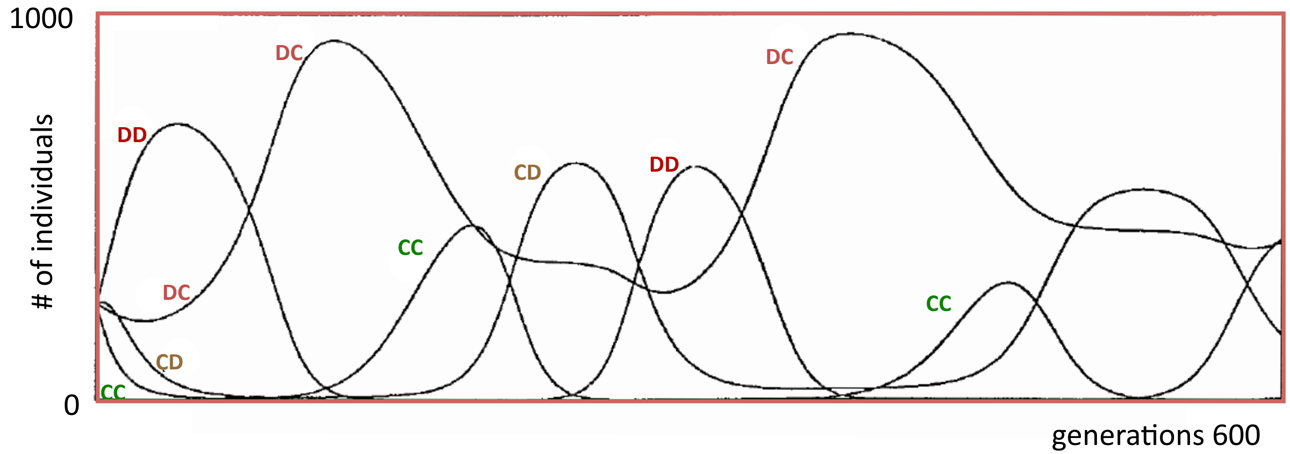
The length L of the code detemines the memory $m = \log_2 L$

Mutations

- Point mutation: **DCDC → DDDC**
- Gene duplication: **DC → DCDC**
- Split mutation (removal): **DDDC → DD**

- Mutation rates ~ 10^{-5}

Evolution simulation: repeated PD with mistakes



Initial population, 250 each of memory 1 strategies:

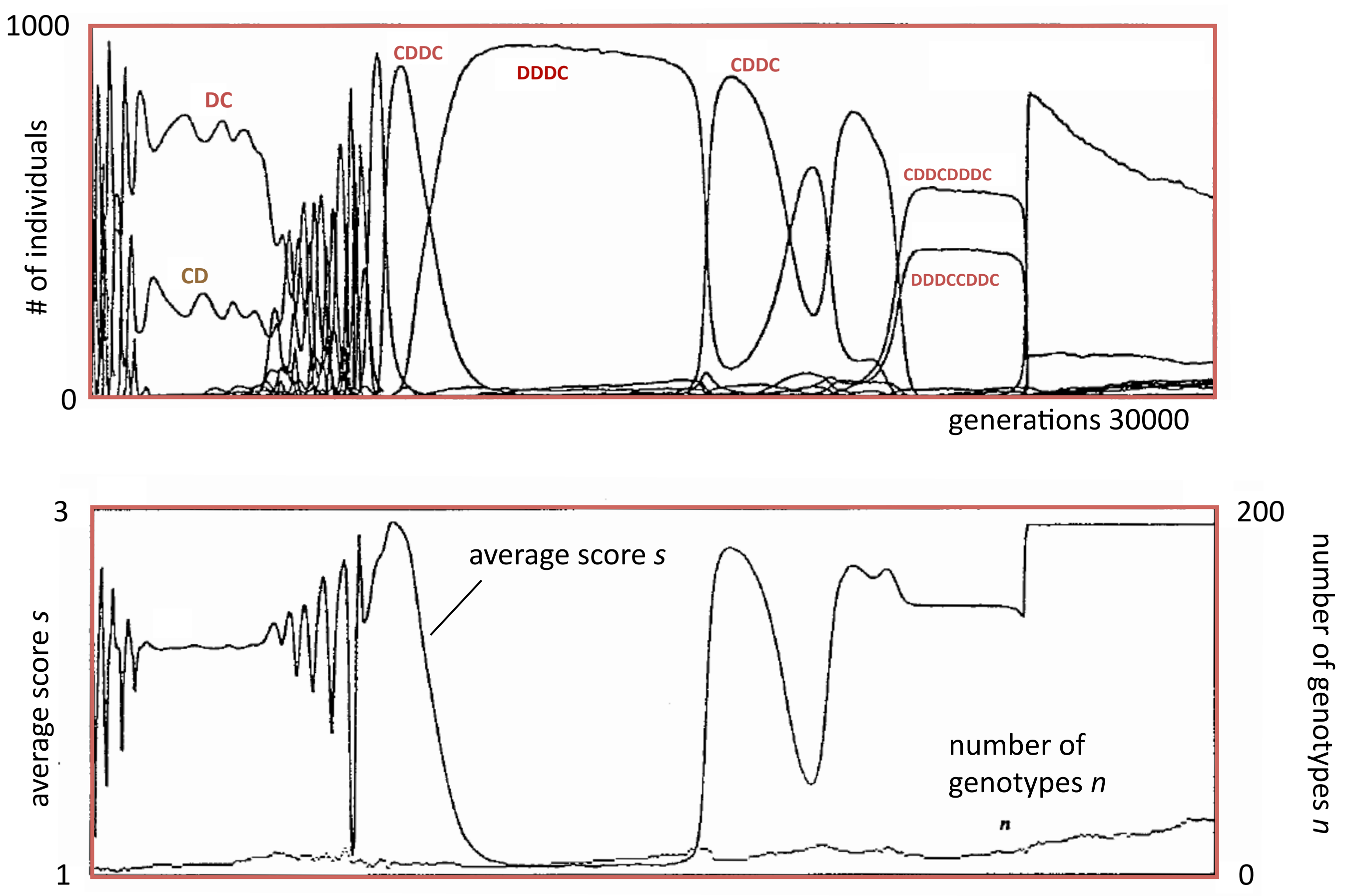
CC — Always Cooperate

DD — Always Defect

DC — Tit-for-tat

CD — Anti-Tit-for-tat

Evolution simulation: repeated PD with mistakes



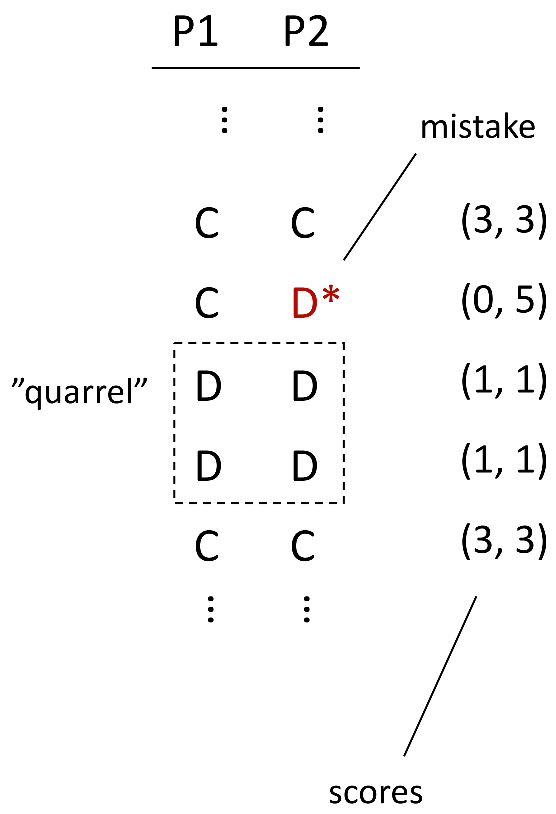
Evolved mechanism for cooperation

Avoiding exploitation: The score gained by defection must be smaller than the cost inflicted by punishment.

If payoffs are

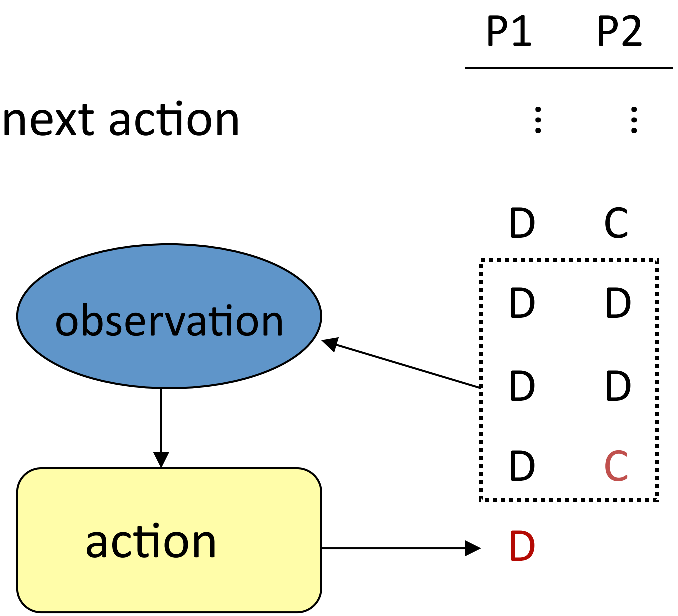
(3, 3)	(0, 5)
(5, 0)	(1, 1)

The player making the mistake gains 5 – 3 = 2 points when defecting, but looses 3 – 1 = 2 points per round during the ”quarrel”, all compared to continued mutual cooperation. Thus the two rounds of quarrel makes it costly to defect by purpose.



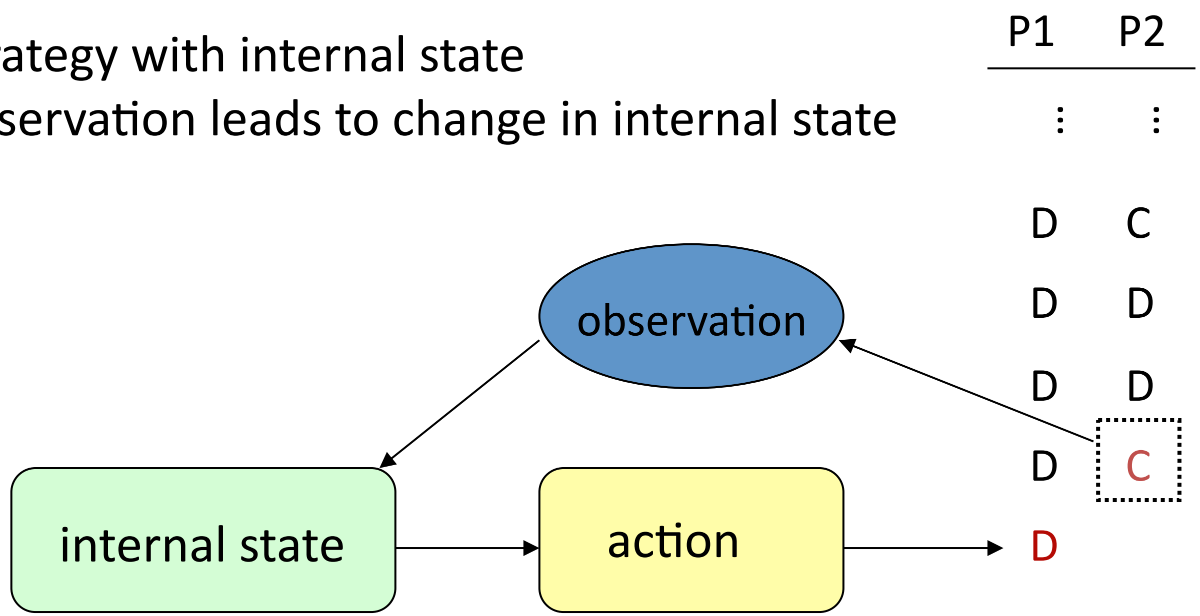
Models of behaviour in repeated games

- Reactive strategy
  mapping a finite history to next action



Models of behaviour in repeated games

- Strategy with internal state
  observation leads to change in internal state

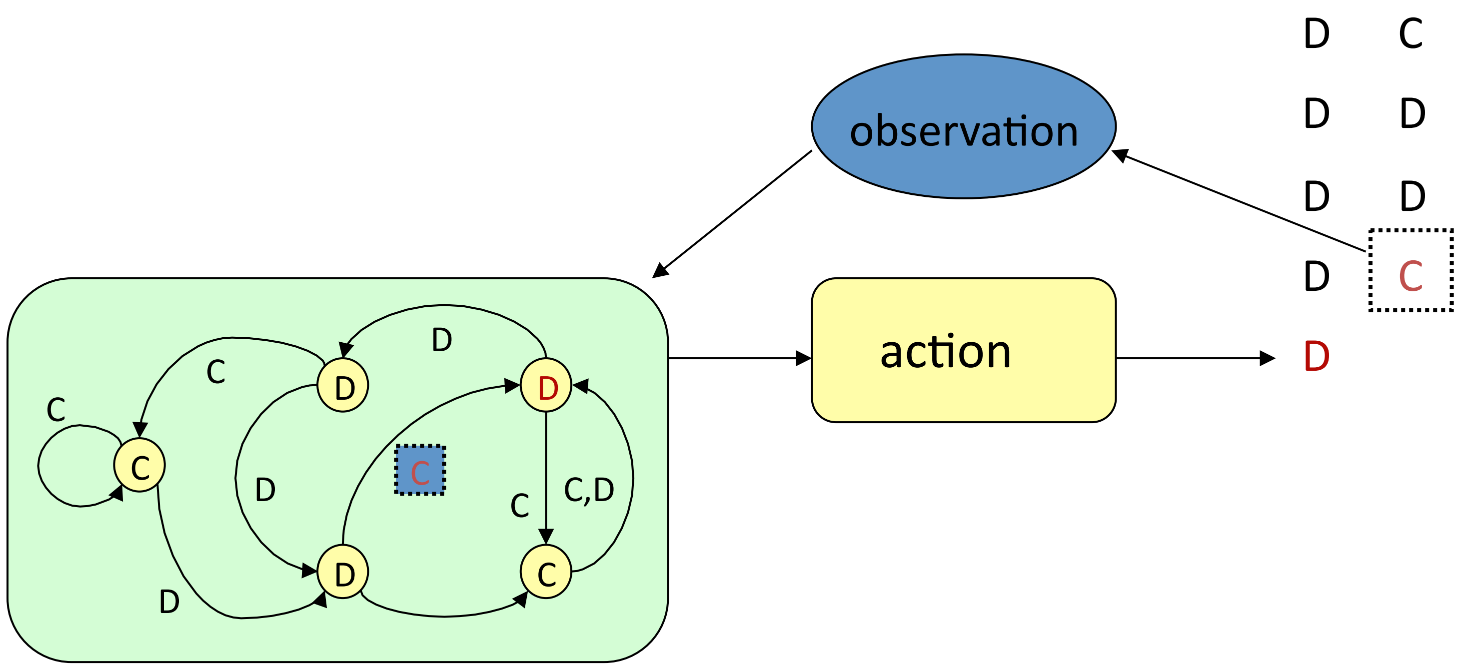


Models of behaviour in repeated games

- Strategy with internal state
  observation leads to change in internal state

P1 P2

⋮ ⋮

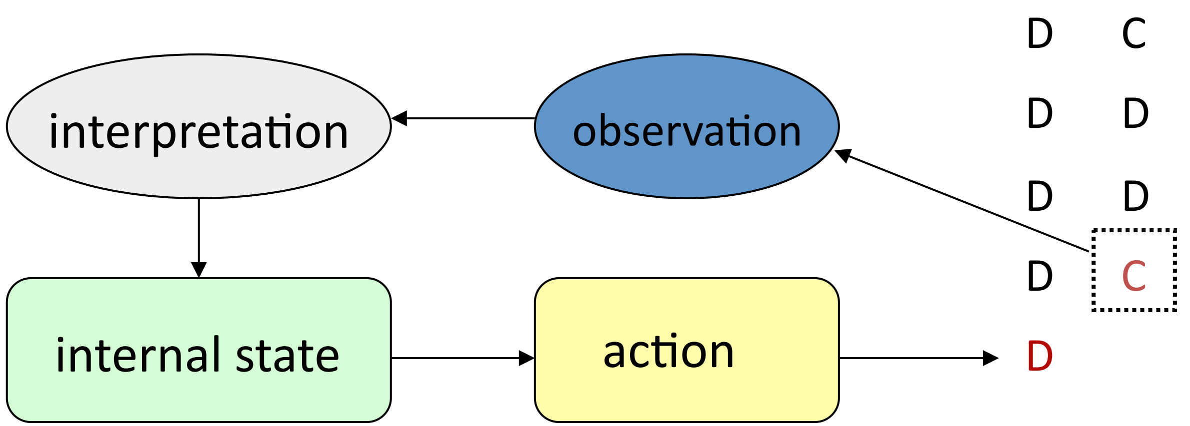


Models of behaviour in repeated games

- Strategy with interpretation
  interpretation of observation → state change

P1 P2

⋮ ⋮



This idea of modelling of the "cognitive processing" in agents have been successfully tested in more complicated games (Eriksson & Lindgren, 2002, 2006).

Nash equilibrium and rationality

- To what extent does the Nash equilibrium (NE) reflect rational choices by the players?
- A Nash equilibrium is often considered an undesirable state when it is not a Pareto optimal solution* (or the Pareto optimal solution with the highest total payoff).
- What mechanisms lead to desirable Pareto optimal states rather than to Nash equilibria (like population structure, evolutionary mechanisms, agent behaviour, etc)?
- What modifications of the game (control mechanisms like constraints, taxes, subsidies, etc) can be applied to change the NE structure?

* Pareto optimal: situation where no one can get better off without reducing the score for another one (using joint action).

Human behaviour and rationality

- Human behaviour in game-like situation exhibits higher degree of cooperation than game-theory suggests: Ultimatum game, Finitely repeated Prisoner's Dilemma.
- Robert Aumann
  - Rule rationality: Actions are based on rules rather than on reasoning. Rules may have evolved and in that way they may seem ”rational”.
  - Act rationality: The actions are chosen on the basis of reasoning about their consequences (traditional rationality).
- Herbert Simon: Bounded rationality — ”boundedly rational agents experience limits in formulating and solving complex problems and in processing … information”