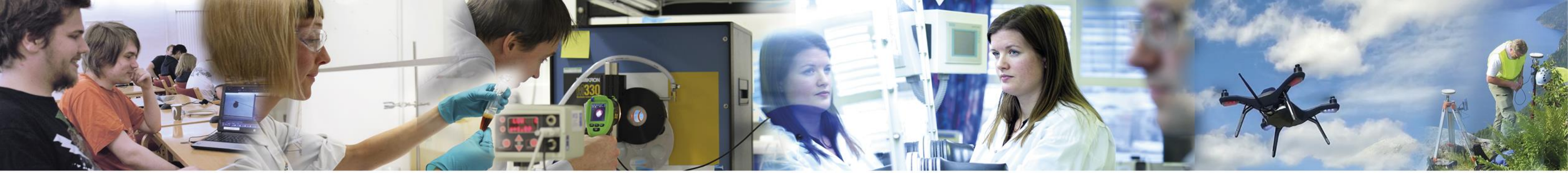


UiO : Department of Technology Systems
University of Oslo

TEK5010/9010 - Multiagent systems 2020
Lecture 8

Non-cooperative game theory

Jonas Moen

Corona restrictions at UiO

Remember to keep everyone safe by:

1. Washing hands
2. Keeping your distance (1 metre)
3. Staying home if you are sick

https://www.uio.no/english/about/hse/corona/index.html

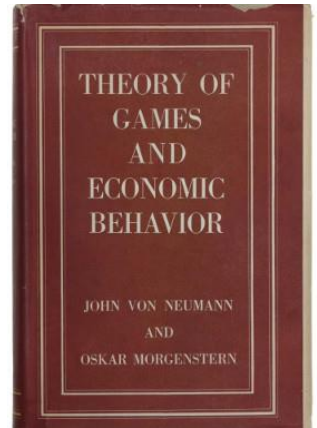
Highlights lecture 8 – Non-cooperative game theory*

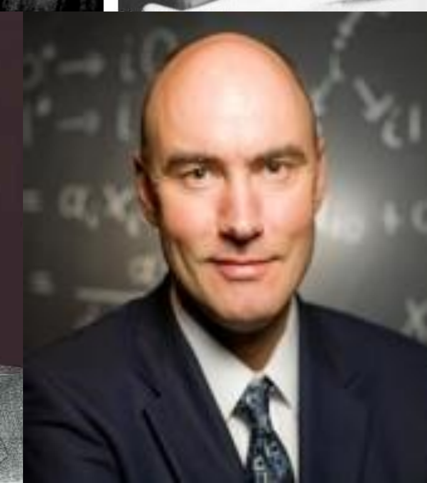
- Classification of game theory
- Utility of self-interested agents
- Strategic interaction and strategic games
- Solution concepts
- Prisoner's dilemma and the iterated PD
- Program equilibria

*Wooldridge, 2009: chapter 11

A quick survey of game theory

- Non-cooperative games:
  self-interested agents
- Cooperative games:
  agents forming coalitions
- Evolutionary games:
  payoffs are frequency dependent
- Behavioural game:
  discrepancy between theory
  and reality





Images: thatsmaths, Harvard, Balzan

Self-interested agents

Agents have their own desires and beliefs

1. Desires are modelled by maximizing expected utility*

$$Ag_{opt} = \max_{Ag \in AG_m} \sum_{r \in R(Ag, Env)}^{k} u(r)P(r|Ag, Env)$$

2. Beliefs are modelled by information processes

*MAS chapter 2, The intelligent agent

Outcomes

$$\Omega = \{\omega_1, \omega_2, \dots\}$$

where Ω is a set of outcomes that agents can have

ω_i is outcome

Utility

$$u_i : \Omega \to \mathbb{R} \qquad (\text{e.g. } u_i(\omega_1) \to \mathbb{R})$$

where u_i is utility of agent i

Ω is the set of possible outcomes

$\mathbb{R}$ is the set of real numbers

ω_1 is a particular outcome

Preference ordering

Agents are able to rank outcomes:

$$u_i(\omega) \geq u_i(\omega') \Leftrightarrow \omega \succcurlyeq_i \omega'$$

meaning agent i prefers outcome ω over ω' or is indifferent

$$u_i(\omega) > u_i(\omega') \Leftrightarrow \omega \succ_i \omega$$

meaning agent i strictly prefers outcome ω over ω'

Properties of the preference ordering

1. Reflexivity
   For all $\omega \in \Omega$, we have that $\omega \succcurlyeq_i \omega$
2. Transitivity
   If $\omega \succcurlyeq_i \omega'$, and $\omega' \succcurlyeq_i \omega''$ then $\omega \succcurlyeq_i \omega''$
3. Comparability
   For all $\omega \in \Omega$, and $\omega' \in \Omega$ we have either $\omega \succcurlyeq_i \omega'$ or $\omega' \succcurlyeq_i \omega$

Utility and money

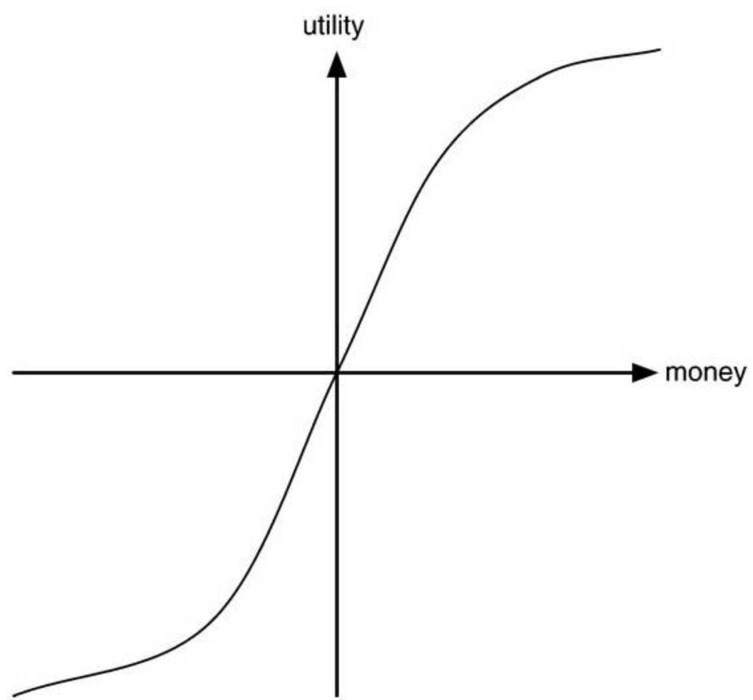


Image: Figure 11.2, Wooldridge 2009

Strategic interaction

Basic idea:

«What I do depend on what you do, and what you do depend on what I do… which we both should have taken into account in the first place.»

Strategic interaction

Basic idea:

The environment is altered in simultaneous actions by agents.

Assume:

1. Agents must act
2. Agents can not see other agents perform actions

Strategic interaction

Mathematically,

$$\tau : Ac_i \times Ac_j \rightarrow \Omega$$

where τ is state transformer function

Ac_i is action of agent i

Ω is the set of outcomes

Strategic interaction

The simplest strategic game conceivable:

2 agents, i and j, with 2 actions available, C and D,

‘C’ for Cooperate

‘D’ for Defect

Strategic interaction

Let us find the possible action combinations:

$$(C, C) \lor (C, D) \lor (D, C) \lor (D, D)$$

Giving 4 possible outcomes:

$$\tau(C, C) = \omega_1$$
$$\tau(C, D) = \omega_2$$
$$\tau(D, C) = \omega_3$$
$$\tau(D, D) = \omega_4$$

Strategic interaction

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

How do agents evaluate these 4 outcomes?

Agent *i:*

$$u_i(\omega_1) = u_{i1} = payoff_{1,i}$$
$$u_i(\omega_2) = u_{i2} = payoff_{2,i}$$
$$u_i(\omega_3) = u_{i3} = payoff_{3,i}$$
$$u_i(\omega_4) = u_{i4} = payoff_{4,i}$$

Agent *j:*

$$u_j(\omega_1) = u_{j1} = payoff_{1,j}$$
$$u_j(\omega_2) = u_{j2} = payoff_{2,j}$$
$$u_j(\omega_3) = u_{j3} = payoff_{3,j}$$
$$u_j(\omega_4) = u_{j4} = payoff_{4,j}$$

Game in strategic form*

Outcome matrix:

i \ j	D	C
D	ω_4	ω_3
C	ω_2	ω_1

*Also called normal form

Game in strategic form*

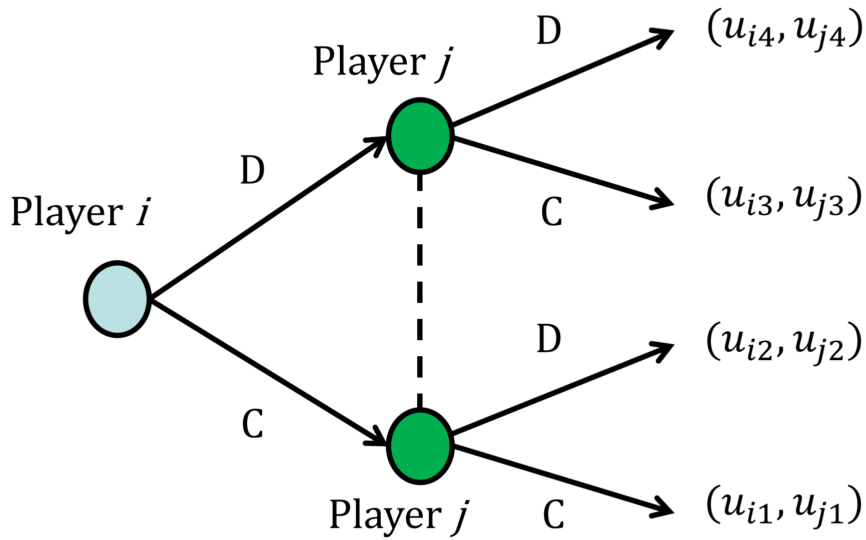
Payoff matrix (in utility):

i \ j	D	C
D	u_{i4}, u_{j4}	u_{i3}, u_{j3}
C	u_{i2}, u_{j2}	u_{i1}, u_{j1}

*Also called normal form

Game in extensive form*

Payoffs (in utility):



*Also called a game tree

Solution concepts

1. Maximizing social welfare
2. Pareto efficiency
3. Dominant strategy
4. Nash equilibrium

Maximizing social walfare

Chose the strategy that gives the highest aggregated utility among all agents.

$$sw(\omega_i) = \sum_{j \in Ag} u_j(\omega_i)$$

where $sw(\omega_i)$ is social welfare of outcome ω_i

u_j is utility for agent j of outcome ω_i

Pareto efficiency

A solution is Pareto efficient if no improvement is possible without making someone else worse off.

Also called Pareto optimality.

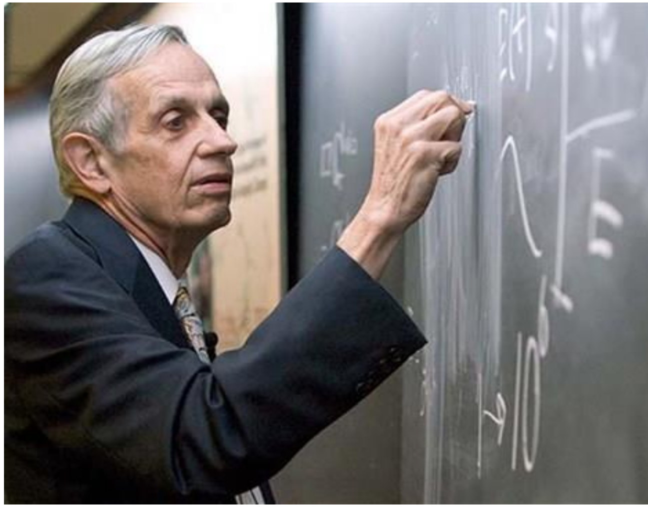
This is a central concept in economics and in multi-objective optimization.

Dominant strategy

A strategy s for agent i is dominant if s it is best respons to all of agent j's strategies s'.

There is no guarantee of the existence of such a solution.

Nash equilibrium



The two strategies s_i and s_j of agents i and j are in Nash equilibrium

1. if player i plays s_i, player j can do no better than playing s_j
2. if player j plays s_j, player i can do no better than playing s_i

s_i and s_j are best response to each other, no player regret their strategy choice.

Image: dreamingtheworld.tv

Nash equilibrium

Two types of Nash equilibria

1. Pure strategy Nash equilibrium
2. Mixed strategy Nash equilibrium

Nash's theorem guarantees the existence of a solution in mixed strategy games only.

Pure strategy Nash equilibrium

Check all combinations of N agents and M strategies

1. This gives a computational complexity of $\mathcal{O}(M^N)$, which is acceptable for small M and N
2. There might not exist a pure strategy Nash equilibrium
3. There might be more than one pure strategy Nash equilibrium

Mixed strategy Nash equilibrium

Include the probability of playing different strategies.

The solution concept becomes to find the optimal probabilities of playing the various strategies. How do you play the game? How often do you play a particular strategy?

A mixed strategy over $(s_1, s_2, \ldots, s_M)$ strategies is to find a probability distribution $(p_1, p_2, \ldots, p_M)$ of playing the different strategies $(s_1, s_2, \ldots, s_M)$.

Nash's theorem

Every game in which every player has a finite set of possibilities has a Nash equilibrium in mixed strategies.

Note:

Often difficult to find Nash equilibrium in mixed strategies due to high computational complexity, but they do exist!

The Prisoner’s dilemma (PD)

The most famous game in game theory

The Prisoner's dilemma

«Two men are collectively charged with a crime and held in separate cells. They have no way of communicating with each other or making any kind of agreement. The two men are told that:

1. *If one of them confesses to the crime and the other does not, the confessor will be freed, and the other will be jailed for 3 years.*
2. *If both confess to the crime, then each will be jailed for 2 years.*

Both prisoners know that if neither confesses, then they will be jailed for 1 year.» [Wooldridge, 2009]

The Prisoner's dilemma

Let us model the game:

1. Who are the players?
2. What are their available strategies?
3. What are the possible outcomes?
4. What are the payoffs (how do the players evaluate the outcomes)?

The Prisoner's dilemma

Let us model the game:

1. Who are the players?

Prisoner i and prisoner j, making it a 2 player game N=2.

The Prisoner's dilemma

Let us model the game:

1. Who are the players? Agent i and j, N=2
2. **What are their available strategies?**

2 possible strategies for each player, either
Cooperate (C) or *Defect* (D), making $S \in \{C, D\}$, M=2.

The Prisoner's dilemma

Let us model the game:

1. Who are the players? Agent i and j, N=2
2. What are their available strategies? $S \in \{C, D\}$, M=2
3. **What are the possible outcomes?**

We could have 4 different outcoms $\left(s_i = Ac_{l,i}, s_j = Ac_{k,j}\right)$:
(C, C), (D, C), (C, D) or (D, D) ⇔ (1,1), (0,3),(3,0) or (2,2) years

The Prisoner's dilemma

Let us model the game:

1. Who are the players? Agent i and j, N=2
2. What are their available strategies? $S \in \{C, D\}$, M=2
3. What are the outcomes? (1,1), (0,3), (3,0) or (2,2) years
4. **What are the payoffs?**

$$\left.\begin{aligned} u_i(0y) &= u_j(0y) = 5 \text{ utility} \\ u_i(1y) &= u_j(1y) = 3 \text{ utility} \\ u_i(2y) &= u_j(2y) = 2 \text{ utility} \\ u_i(3y) &= u_j(3y) = 0 \text{ utility} \end{aligned}\right\} (3,3),\ (5,0),\ (0,5) \text{ or } (2,2)$$

The Prisoner's dilemma

Let us model the game:

1. Who are the players? **Agent i and j, N=2**
2. What are their available strategies? $S \in \{C, D\}$, M=2
3. What are the outcomes? (C, C), (D, C), (C, D) **or** (D, D)
4. What are the payoffs? **(3,3), (5,0), (0,5) or (2,2) utility**

⇒ Symmetric 2×2 interaction on strategic form

The Prisoner's dilemma

Payoff matrix:

i \ j	**D**	**C**
D	2,2	5,0
C	0,5	3,3

The Prisoner's dilemma

Maximizing social welfare:

i \ j	D	C
D	2,2 (2+2=4)	5,0 (5+0=5)
C	0,5 (0+5=5)	3,3 (3+3=6) *

Chose the strategy that gives the highest aggregated utility among all agents.

The Prisoner's dilemma

Pareto efficiency:

i \ j	D	C
D	2,2	5,0 *
C	0,5 *	3,3 *

A solution is Pareto efficient if no improvement is possible without making someone else worse off.

The Prisoner's dilemma

Dominant strategy:

i \ j	D	C
D	2,2 *	5,0
C	0,5	3,3

A stragegy s for agent i is dominant if s it is best respons to all of agent j's strategies s'.

The Prisoner's dilemma

Nash equilibrium:

i \ j	D	C
D	2,2 *	5,0
C	0,5	3,3

s_i and s_j are best respons to each other, no player regret their strategy choice. Check all combinations of N agents and M strategies.

The Prisoner's dilemma

Why is it called a dilemma?

Solution concept	Solution	Payoffs	Social welfare, $\sum u$
Maximizing social welfare	(C,C)	(3,3)	6
Pareto efficiency	(C,C), (D,C), (C,D)	(3,3), (5,0), (0,5)	6, 5, 5
Dominant strategy	(D,D)	(2,2)	4
Nash equilibrium	(D,D)	(2,2)	4

The notion that rational agents could do better by cooperating.

The Prisoner's dilemma

Important real-world game:

- «Tragedy of the commons», [Hardin, 1968]
  - Grazing livestock
  - Overfishing the seas
  - Capacity bandwidth on the Internet
- Nuclear weapons treaties
- What is cooperation in biology?

The Prisoner's dilemma

Can we have cooperation and rationality at the same time? [Binmore, 1992]

- Are we altruists? Affects the payoffs, not PD anymore.
- How about including punishment? Also not PD anymore.
- Group selection and kin selection? Selfish genes?
- People are not rational for small utilities, but in life and death situations we prefer the the rational outcome.

Iterated Prisoner's dilemma (IPD)

By repeating the Prisoner's dilemma over many rounds the chance of cooperation increases, mainly due to:

- The threat of «punishment» by defecting in subsequent rounds
- Loss of utility can be «amortized» over many rounds

Iterated Prisoner's dilemma

How does repeating the game affect the outcomes?

1. Infinite rounds of PD
   Cooperation is rational outcome due to threat of defection.
2. Fixed number of rounds PD
   Rational to defect in last round, i.e. 'backward induction'.
3. Non-zero probability of future PD round
   Rational to cooperate if probability of one more round is large enough compared to the payoffs.

Axelrod's tournament

«The Evolution of Cooperation», [Axelrod, 1984].

- Best-known piece of multiagent system reserach.
- How can cooperation arise in societies of self-interested agents?
- Tested different submitted strategies for the iterated PD.
- Winner was best overall strategy against all other strategies tested in 200 rounds of IPD.

Image: Youtube

Axelrod's tournament

Some strategies submitted:

- *Random*, 50/50 C or D
- *All-D*; only D
- *Tit-For-Tat* (TFT); first C then repeat opponent
- *Tester*; first D then change to TFT is opponent D
- *Joss*; TFT but 10% D
- …

Axelrod's tournament

Overall winner was TFT…

…but TFT will lose to *All-D*.

Axelrod's tournament

Rules for success in iterated PD

- Do not be envious, don't try to beat opponent.
- Do not be first to defect, instead amortize loss
- Reciprocate C and D, balanced forgiveness and retaliation is necessary
- Do not be too clever, TFT was simplest strategy
  - Too complex for opponent to understand, appear random
  - Overgeneralization of opponents model

Program equilibria

Basic idea is to compare strategies before conditional action is taken by a moderator.

«I will cooperate if you will»

Proposed by [Tennenholtz, 2004] and is subject of much ongoing research.

Program equilibria

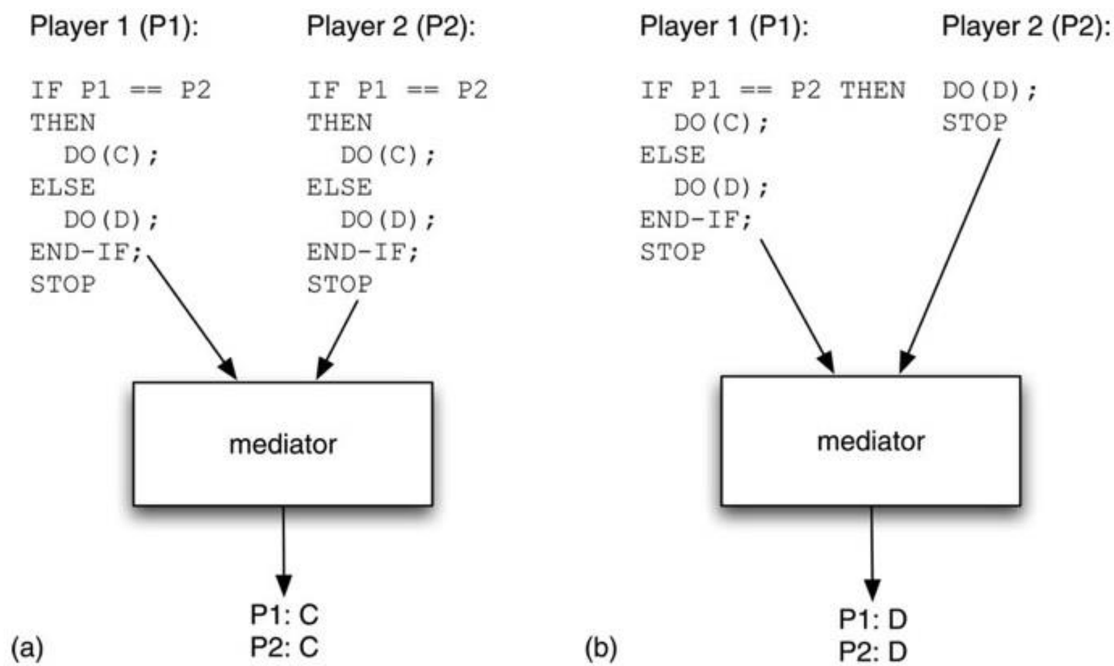
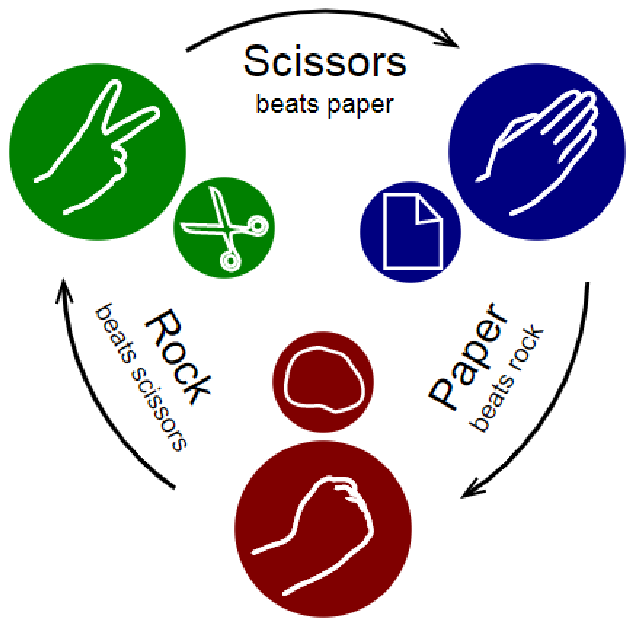


Image: Figure 11.3, Wooldridge 2009

Rock-paper-scissors game



How to win this game?

Image: Wikipedia

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

Payoffs in normal form

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

Nash?

Rock-paper-scissors game

j / i	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

No pure strategy Nash!

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

Mixed strategy Nash?

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

$E_j(U_R)$ $E_j(U_P)$ $E_j(U_S)$

Player *j* must chose between strategy {R,P,S},

i.e. maximizing expected utility $\max\{E_j(U_R), E_j(U_P), E_j(U_S)\}$

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

$E_j(U_R)$

Expected utility of player *j* chosing *Rock* depends on player *i*

$$E_j(U_R)=p_{iR}U_{jRR}+p_{iP}U_{jPR}+p_{iS}U_{jSR}$$

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)

$E_j(U_R)$ $E_j(U_P)$ $E_j(U_S)$

General idea: Player *j* is not able to change outcome if player *i* plays a mixed strategy of p_{iR}, p_{iP}, p_{iS} giving $E_j(U_R)=E_j(U_P)=E_j(U_S)=U$

Rock-paper-scissors game

i \ j	R	P	S
R	(0,0)	(-1,1)	(1,-1)
P	(1,-1)	(0,0)	(-1,1)
S	(-1,1)	(1,-1)	(0,0)
	U	U	U

Why? Expected utility of player *j* is given by $p_{jR}U+p_{jP}U+p_{jS}U=(p_{jR}+p_{jP}+p_{jS})U=U$ since $(p_{jR}+p_{jP}+p_{jS})=1$ and assuming $E_j(U_R)=E_j(U_P)=E_j(U_S)=U$

Rock-paper-scissors game

i \ j	R	P	S	
R	(0,0)	(-1,1)	(1,-1)	U_i
P	(1,-1)	(0,0)	(-1,1)	U_i
S	(-1,1)	(1,-1)	(0,0)	U_i
	U_j	U_j	U_j	

And vice versa! Meaning that none of the players have any reason for changing their mixed strategy, i.e. Nash equilibrium!

Rock-paper-scissors game

i \ j	R	P	S	
R	(0,0)	(-1,1)	(1,-1)	U_i
P	(1,-1)	(0,0)	(-1,1)	U_i
S	(-1,1)	(1,-1)	(0,0)	U_i
	U_j	U_j	U_j	

For player *i*:

Eq 1: $p_{iR}0 + p_{iP}(-1) + p_{iS}1 = p_{iS} - p_{iP}$

Eq 2: $p_{iR}1 + p_{iP}0 + p_{iS}(-1) = p_{iR} - p_{iS}$

Eq 3: $p_{iR}(-1) + p_{iP}1 + p_{iS}0 = p_{iP} - p_{iR}$

Rock-paper-scissors game

i \ j	R	P	S	
R	(0,0)	(-1,1)	(1,-1)	U_i
P	(1,-1)	(0,0)	(-1,1)	U_i
S	(-1,1)	(1,-1)	(0,0)	U_i
	U_j	U_j	U_j	

For player *i*:

Eq 1 and eq 2: $p_{iS} - p_{iP} = p_{iR} - p_{iS} \Longrightarrow p_{iR} = 2p_{iS} - p_{iP}$

And : $p_{iS} + p_{iP} + p_{iS} = 1 \Longrightarrow p_{iP} = 1 - p_{iS} - p_{iR}$

Rock-paper-scissors game

i \ j	R	P	S	
R	(0,0)	(-1,1)	(1,-1)	U_i
P	(1,-1)	(0,0)	(-1,1)	U_i
S	(-1,1)	(1,-1)	(0,0)	U_i
	U_j	U_j	U_j	

For player *i*:

$$p_{iR} = 2p_{iS} - (1 - p_{iS} - p_{iR})$$

$$p_{iS} = 1/3$$

and by symmetry $p_{iR} = p_{iP} = p_{jR} = p_{jP} = p_{jS} = 1/3$

Other symmetric 2x2 games

1. $(C,C) \succ_i (C,D) \succ_i (D,C) \succ_i (D,D)$ cooperation dominates
2. $(C,C) \succ_i (C,D) \succ_i (D,D) \succ_i (D,C)$ cooperation dominates
3. $(C,C) \succ_i (D,C) \succ_i (C,D) \succ_i (D,D)$
4. $(C,C) \succ_i (D,C) \succ_i (D,D) \succ_i (C,D)$ stag hunt
5. $(C,C) \succ_i (D,D) \succ_i (C,D) \succ_i (D,C)$
6. $(C,C) \succ_i (C,D) \succ_i (D,C) \succ_i (C,D)$
7. $(C,D) \succ_i (C,C) \succ_i (D,C) \succ_i (D,D)$
8. $(C,D) \succ_i (C,C) \succ_i (D,D) \succ_i (D,C)$
9. $(C,D) \succ_i (D,C) \succ_i (C,C) \succ_i (D,D)$
10. $(C,D) \succ_i (D,C) \succ_i (D,D) \succ_i (C,C)$
11. $(C,D) \succ_i (D,D) \succ_i (C,C) \succ_i (D,C)$
12. $(C,D) \succ_i (D,D) \succ_i (D,C) \succ_i (C,C)$

Image: Table 11.1, Wooldridge 2009

Other symmetric 2x2 games

13. $(D, C) \succ_i (C, C) \succ_i (C, D) \succ_i (D, D)$ game of chicken
14. $(D, C) \succ_i (C, C) \succ_i (D, D) \succ_i (C, D)$ prisoner's dilemma
15. $(D, C) \succ_i (C, D) \succ_i (C, C) \succ_i (D, D)$
16. $(D, C) \succ_i (C, D) \succ_i (C, C) \succ_i (C, C)$
17. $(D, C) \succ_i (D, D) \succ_i (C, C) \succ_i (C, D)$
18. $(D, C) \succ_i (D, D) \succ_i (C, D) \succ_i (C, C)$
19. $(D, D) \succ_i (C, C) \succ_i (C, D) \succ_i (D, C)$
20. $(D, D) \succ_i (C, C) \succ_i (D, C) \succ_i (C, D)$
21. $(D, D) \succ_i (C, D) \succ_i (C, C) \succ_i (D, C)$
22. $(D, D) \succ_i (C, D) \succ_i (D, C) \succ_i (C, C)$
23. $(D, D) \succ_i (D, C) \succ_i (C, C) \succ_i (C, D)$ defection dominates
24. $(D, D) \succ_i (D, C) \succ_i (C, D) \succ_i (C, C)$ defection dominates

Image: Table 11.1, Wooldridge 2009

The stag hunt

Payoff matrix:

i \ j	D	C
D	1,1 *	2,0
C	0,2	3,3 *

You and a friend plan to apperar with ridiculous haircut on last school day. [Rousseau, 1775]

Game of chicken

Payoff matrix:

i \ *j*	D	C
D	0,0	3,1 *
C	1,3 *	2,2

You and an opponent drive cars toward the edge of a cliff, first to turn is a chicken. D is drive, C is turn.

Competitive interactions

An iteraction is said to be strictly competitive among agent i and agent j when

$$\omega \succ_i \omega' \text{ if and only if } \omega' \succ_j \omega$$

for outcome ω and ω'.

Zero-sum interactions

Zero-sum games are formally described as

$$u_i(\omega) + u_j(\omega) = 0 \text{ for all } \omega \in \Omega$$

where $u_i(\omega)$ is utility of agent i of outcome ω

Relation to real-world applications is questionable.
[Zagare,1984]

In conclusion

Non-cooperative game theory raises the question of «what is cooperation?» in biology, sociology, economics, computer science…

- How does cooperation emerge?
- How is cooperation maintained?

… under the threat of opportunism.

Summary lecture 8 – Non-cooperative game theory*

- Classification of game theory
- Utility of self-interested agents
- Strategic interaction and strategic games
- Solution concepts (SW, PE, DS, Nash pure and mixed)
- Prisoner's dilemma and the iterated PD
- Program equilibria
- What is cooperation?

*Wooldridge, 2009: chapter 11