Artificial Intelligence 2

Roman Barták
Department of Theoretical Computer Science and Mathematical Logic

Decisions with multiple agents

What if the uncertainty is due to other agents and the decisions they make? And what if the decisions of those agents are in turn influenced by our decisions?

- **agent design**
  - **game theory** can analyze the agent's decisions and compute the expected utility for each decision (under the assumption that other agents are acting optimally according to game theory)
- **mechanism design**
  - **inverse game theory** make it possible to define the rules of the environment so that the collective good of all agents is maximized (when each agent adopts the game-theoretic solution that maximizes its own utility)

Single-move games

Consider a restricted set of games, where all players take action simultaneously and the result of the game is based on this single set of actions

- what matters is that no player has knowledge of the other players' choices

A single-move game is defined by three components:

- **players** (or agents), like O (odd) and E (even)
- **actions** that the players can choose, like one or two fingers
- **a payoff function** that gives the utility to each player for each combination of actions by all players; the payoff matrix for two-finger Morra is as follows:

	O: one	**O: two**
E: one	E=+2, O=-2	E=-3, O=+3
E: two	E=-3, O=+3	E=+4, O=-4

Single-move games: solution and strategy

Each player in a game must adopt and then execute a **strategy** (policy).

- **a pure strategy**

  is a deterministic policy; for a single-move game, it is just a single action
- **a mixed strategy**

  is randomized policy that selects actions according to a probability distribution; for two actions, it is written [p,a; (1-p),b]

The game's outcome is a numeric value for each player.

A solution to a game is a strategy profile (an assignment of strategy to each player) in which each player adopts a rational strategy.

- What does "rational" mean when each agent chooses only part of the strategy profile that determines the outcome?

Prisoner's dilemma

Consider the following story:

- Two alleged burglars, Alice and Bob, are caught red-handed near the scene of burglary and are interrogated separately.
- A prosecutor offers each a deal: if you testify against your partner as the leader of a burglary ring, you will go free while your partner will serve 10 years in prison.
- However, if you both testify against each other, you will both get 5 years.
- If you both refuse to testify, you will serve only 1 year each for lesser charge of possessing stole property.

Show they testify or refuse?

- The rational decision is to testify.

	Alice: testify	**Alice: refuse**
Bob: testify	A=-5, B=-5	A=-10, B=0
Bob: refuse	A=0, B=-10	A=-1, B=-1

Dominance

Testify is a **dominant strategy** for the Prisoner's dilemma.

- a strategy s for player p **strongly dominates** strategy s' if the outcome for s is better for p than the outcome for s', for every choice of strategies by the other player(s)
- a strategy s **weakly dominates** s' if s is better on at least one strategy profile and no worse on any other

It is **irrational** to play a dominated strategy and not to play a dominant strategy if one exists.

When each player has a dominant strategy, the combination of those strategies is called a **dominant strategy equilibrium**.

A strategy profile forms an **equilibrium** if no player can benefit by switching strategies, given that every other player sticks with the same strategy. Every game has at least one equilibrium – **Nash equilibrium**.

The outcome of game is **Pareto optimal** if there is no other outcome that all players would prefer.

- An outcome is **Pareto dominated** by another outcome if all players would prefer the other outcome.

Prisoner's dilemma is due to having a dominant strategy equilibrium (testify, testify) that is Pareto dominated by outcome (refuse, refuse).

No dominant strategy

Consider the following game

- Acme, a video game console manufacturer, has to decide whether its next game machine will use Blu-ray discs or DVDs.
- Meanwhile, the video game software producer Best needs to decide whether to produce next game on Blu-ray or DVD.
- The profits of both will be positive if they agree and negative if they disagree.

	Acme: bluray	Acme: dvd
Best: bluray	A=+9, B=+9	A=-4, B=-1
Best: dvd	A=-3, B=-1	A=+5, B=+5

There is **no dominant strategy equilibrium** for this game, but there are **two Nash equilibria**.

There are multiple acceptable solutions, but if each agent aims for a different one, then both agents will suffer.

How can they agree on a solution?

Both can should **choose the Pareto-optimal Nash equilibrium** provided that one exists; (bluray, bluray) is the Pareto-optimal solution

What if there are much such solutions (for example if (bluray, bluray) had payoff (5,5))?

- agents can either guess or communicate
- **coordination games** (games in which players need to communicate)

Mixed strategies

Consider two-finger Morra game

- no pure-strategy profile exists
  - if the total number of fingers is even, then O will want to switch
  - if the total is odd, then E will want to switch
- we must look for **mixed strategies** instead

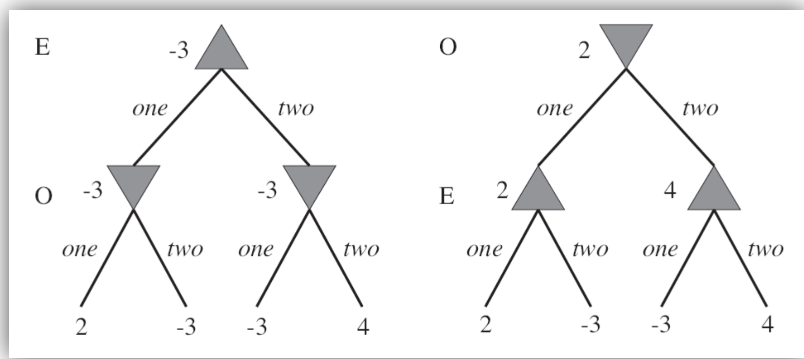
Von Neumann developed a method for finding the optimal mixed strategy for **two-player, zero-sum games** (games in which the sum of the payoffs is always zero).

- the maximin technique
- we need to consider the payoffs of only one player(E)

Maximin technique (pure strategies)

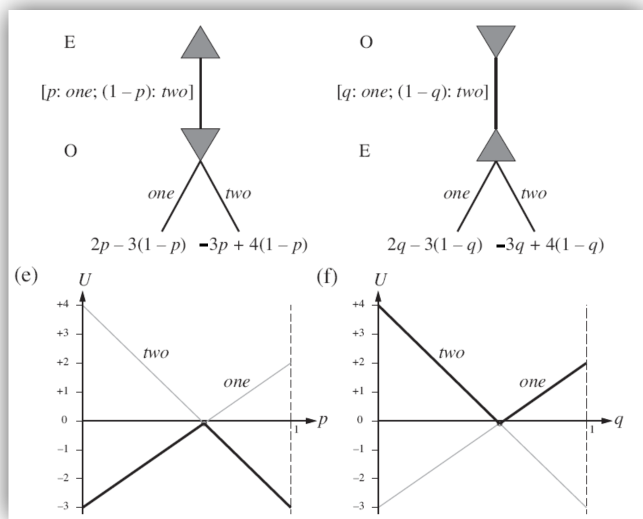
Suppose we change the rules as follows:

- First E picks her strategy and reveals it to O. Then O picks his strategy, with knowledge of E's strategy.
  - This gives a turn-taking game to which we can apply the standard minimax algorithm.
  - Clearly, this games favors O, so we get a **lower bound** for the true utility for E (-3).
- Now suppose we change the rules to force O to reveal his strategy first, followed by E.
  - This gives an **upper bound** for the true utility of E (2).



Maximin technique (mixed strategies)

We need to turn out our analysis to mixed strategies [p,one; (1-p),two]



Once the first player has revealed his or her strategy, the second player might as well choose a pure strategy.

What is the value of p to get the best utility for E (left)?

- p=7/12 and the payoff -1/12

What is the value of q to get the best utility for O (right)?

- q=7/12 and the payoff -1/12

Optimal strategy for both players is [7/12,one; 5/12,two]

- » maximin equilibrium (it is also a Nash equilibrium)
- » two-finger Morra game favor the player O

What if the same game is repeated more times?

Repeated game is the simplest kind of a multiple-move game:

- players face the same choice repeatedly, but each time with knowledge of the history all players' previous choices
- payoffs are additive over time

The repeated version of the prisoner's dilemma:

1. the same players play 100 rounds
   - rational strategy is still to testify (the last game is not the repeated game etc.)
   - earning a total jail sentence of 500 years each
2. 99% chance that the players meet again
   - the expected number of rounds is still 100, but neither player knows for sure which round will be the last
   - **perpetual punishment** strategy: each player *refuses* unless the other play has ever played *testify*
   - the expected future payoff is -100 ($\sum_{t=0}^{\infty} 0.99^t * (-1)$) if both players adopted this strategy
   - a player who deviates from the strategy and chooses *testify* will gain a score 0, but then both players will play testify and the total expected future payoff becomes -495 ($0 + \sum_{t=1}^{\infty} 0.99^t * (-5)$)

A famous strategy is called **tit-for-tat**:

- starting with *refuse* and the echoing the other player's previous move on all subsequent moves
- highly robust and effective against a wide variety of strategies

Mechanism design

So far we focused on the question *„Given a game, what is a rational strategy?“*

What if we ask *„Given that agents pick rational strategies, what game should we design?“*

We would like to design a game whose solutions, consisting of each agent pursuing its own rational strategy, result in the maximization of some global utility function.

This is called **mechanism design** or sometimes **inverse game theory**.

It is used in economics and political science. In general it allows us to construct smart systems out of collection of more limited (even uncooperative) systems.

A **mechanism** consists of:

- a **language** for describing the set of allowable strategies that agents may adopt,
- a distinguished agent – **center** – that collects reports of strategy choices from agents in the games
- an **outcome rule**, known to all agents, that the center uses to determine the payoffs of each agent given their strategy choices

Auctions

An **auction** is a mechanism for selling some goods to members of a pool of bidders.

For simplicity, we concentrate on auctions with a single item for sale.

Each bidder i has a utility value v_i for having the item.

- In some cases, each bidder has a **private value** for the item.
  - An old furniture has different value for a furniture collector and young family.
- In other cases, the item has a **common value**, but there is uncertainty as to what the actual value is.
  - Different bidders have different information and hence different estimates of the item’s true value.

Auction mechanism

- each bidder gets a chance to make a bid b_i
- the highest bid b_{max} wins the item, but the price paid need not be b_{max} (part of mechanism design)

English auction

The best-known auction mechanism is the **ascending-bid**, or **English auction**.

- The center starts by asking for a minimum (or **reserve**) bid b_{min}
- If some bidder is willing to pay that amount, the center then asks for b_{min} + d, for some increment d, and continues up from there.
- The auction ends when nobody is willing to bid anymore.
- Then the last bidder wins the item, paying the price he bid.

How do we know if this is a good mechanism?

- one goal is to maximize expected revenue for the seller; another goal is to maximize a notion of global utility
- we say an **action is efficient** if the goods go to the agent who values them most

The English auction is usually both efficient and revenue maximizing if there is

- a **sufficient number of bidders** to enter the game
- **no collusion** – an unfair or illegal agreement by two ore more bidders to manipulate prices

Collusion

An unfair or illegal agreement by two ore more bidders to manipulate prices.

It can happen in secret backroom deals or tacitly, within the rules of the mechanism

Example of price manipulation within the rules of the mechanism

- In 1999, Germany auctioned ten blocks of cell-phone spectrum with a simultaneous action (bids were taken on all ten blocks at the same time) using the rule that any bid must be a minimum of a 10% raise over the previous bid on a block.
- There were only two credible bidders, Mannesman and T-Mobile
- Mannesman entered the bid of 20 million DEM on blocks 1-5 and 18.18 million DEM on blocks 6-10.
- T-Mobile interpreted Mannesman's first bid as an offer: both parties could compute that a 10% raise on 18.18 M is 19.99M. Mannesman's bid was interpreted as an offer "we can get each half of blocks for 20M"

What to do with it?

- a higher reserve price
- a sealed-bid first-price auction
- bring a third bidder

Truth-revealing

In general, both the seller and the global utility function benefit if there are more bidders.

One way to encourage more bidders is to make the the mechanism easier for them.

It is desirable that the bidders have a **dominant strategy,** strategy that works against all other strategies.

– an agent with a dominant strategy can just bid, without wasting time contemplating the other agents' possible strategies

Usually such a strategy involves the bidders revealing their truth value v_i – then it is called a **truth-revealing**, or **truthful**, **auction**.

Properties of English auction

The English auction has most of the **desirable properties**:

– bidders have a **simple dominant strategy**: keep bidding as long as the current cost is below your v_i

– this is not quite truth-revealing, because the winning bidder reveals only that his $v_i \geq b_0 + d$ (we know only a lower bound on v_i)

Some **disadvantages** of the English auction:

– if there is one clearly stronger bidder such that he can always bid higher than any other bidder then the competitors may not enter at all, and the strong bidder ends up winning at the reserve price (**discourage competition**)

– **high communication costs** as the auction takes place in one room or all bidders have to have high-speed, secure communication lines

Sealed-bid auction

An alternative mechanism is the **sealed-bid auction**.

- each bidder makes a single bid and communicates it to the auctioneer without the other bidders seeing it
- the highest bid wins

There is **no** longer a **simple dominant strategy**

- the bid depends on expected bids of other agents agents
- let v_i be your utility value and b_0 be the expected maximum of all the other agents' bids
- then you should bid $b_0+\varepsilon$ (for some small ε), if that is less than v_i

Note that the agent with the highest v_i might not win the auction, reducing the bias toward an advantaged bidder (the auction is more competitive).

Sealed-bid second-price auction

A small change in the mechanism for sealed-bid auctions produces the **sealed-bid second-price auction**, also known as a **Vickrey auction**.

- The winner pays the price of the second-highest bid, b_0, rather than paying his own bid.
- The **dominant strategy** is now simply to bid v_i; the mechanism is truth-revealing.

Why is this a dominant strategy?

the utility of agent i in terms of his bid b_i, his value v_i, and the best bid among the other agents b_0 :

$(v_i - b_0)$ if $b_i > b_0$, otherwise 0

- when $(v_i - b_0) > 0$, then any bid that wins the auction is optimal, and bidding v_i in particular wins the auction
- when $(v_i - b_0) < 0$, than any bid that loses the auction is optimal, and bidding , v_i in particular loses the auction
- so bidding v_i is optimal for all possible values of b_0, and in fact, v_i is the only bid that has this property

Consider another type of game, in which countries set their policy for controlling air pollution.

Each country has a choice

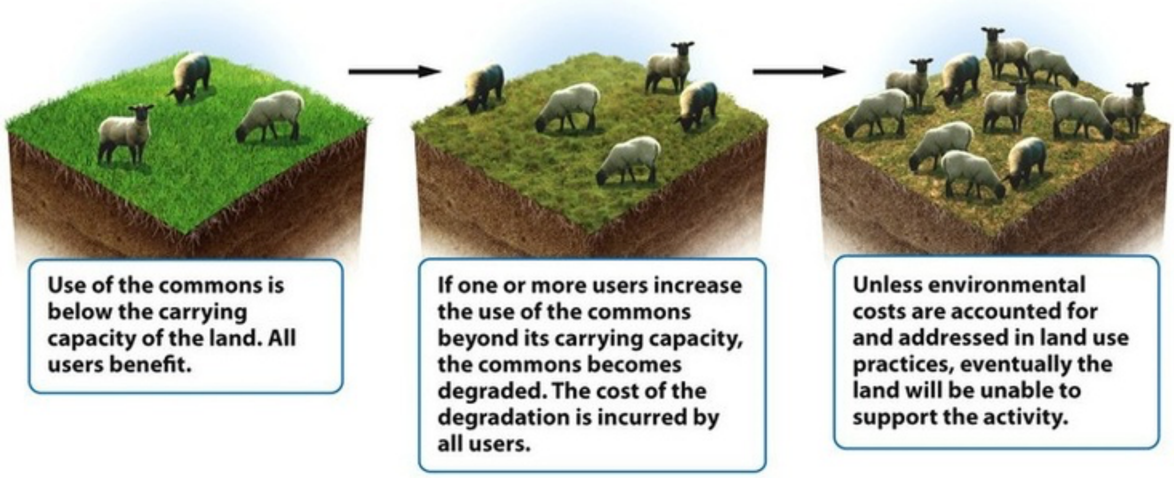
- they can **reduce pollution** at a cost of **-10** points for implementing the necessary changes
- or they can **continue to pollute**, which gives them a net utility of **-5** (in added health costs, etc.) and also contributes **-1** points to every other country (because the air is shared across countries)

What is the strategy of each country?

- **Clearly, the dominant strategy for each country is "continue to pollute".**
- If there are 100 countries and each follows this policy, then each country gets a total utility -104.
- If every country reduces pollution, they would each have a utility of -10!

Tragedy of commons: if nobody has to pay for using a common resource, then it tends to be exploited in a way that leads to a lower total utility for all agents.

It is similar to the prisoner's dilemma: there is another solution to the game that is better for all parties, but there appears to be no way for rational agents to arrive at that solution.



Tragedy of commons: taxes

A standard approach for dealing with the tragedy of commons is to change the mechanism to one that charges each agent for using the commons (a carbon tax).

We need to ensure that all **externalities** – effects on global utility that are not recognized in the individual agents' transactions – are made explicit.

Another example:

- Suppose a city decides it wants to install some free wireless Internet transceivers. However, the number of transceivers they can afford is less than the number of neighborhoods that want them.
- The problem is that if they just ask each neighborhood council "how much do you value this free gift?" they would all have an incentive to lie, and report a high value.
- A solution is asking to pay for it.

Vickrey-Clarks-Groves mechanism

1. the center asks **each agent to report its value** for receiving an item – b_i
2. the **center allocates the goods** to a subset A of the bidders. Let $b_i(A) = b_i$, if $i \in A$, otherwise 0. The center chooses A to maximize total reported utility $B = \Sigma_i b_i(A)$
3. each **agent pays a tax** equal to $W_{-i} - B_{-i}$, where
   $B_{-i} = \Sigma_{j \neq i} b_j(A)$
   W_{-i} = total global utility if i were not in the game
   each winner would pay a tax equal to the highest reported value among the losers (losers pay nothing)

Properties of Vickrey-Clarks-Groves mechanism

Why does the VCG mechanisms make the agents happy?

- all winners should be happy because they pay a tax that is less than their value
- all losers are as happy as they can be, because they value the goods less than the required tax

Why is it that **this mechanism is truth-revealing**?

- each agent maximizes his payoff, which is the value of getting an item, minus the tax

  $v_i(A) - (W_{-i} - B_{-i},)$
- agent i knows that the center will maximize global utility using the reported values

  $\Sigma_j b_j(A) = b_i(A) + \Sigma_{j \neq i} b_j(A)$
- whereas agent i wants the center to maximize

  $v_i(A) + \Sigma_{j \neq i} b_j(A) - W_{-i}$
- Since agent i cannot affects the value of W_{-i} (it depends only on the other agents), the only way i can make the center optimize what i wants is to report the true utility

  $b_i = v_i$

Department of Theoretical Computer Science and Mathematical Logic
bartak@ktiml.mff.cuni.cz