01DRO1
Decision Making under Uncertainty

2017/2018

Multi-agent systems.
Elements of Game Theory.

Lecture 11

22.5.2018

Readings:

Y.Shoham, K. Leyton-Brown (2008) *Multiagent Systems Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press.

Osborne, M. J. (2004). *An Introduction to Game Theory*. Oxford University Press.

Claus, C., Boutilier, C. (1998). *The dynamics of reinforcement learning in cooperative multiagent systems*. In AAAI/IAAI, pages 746- 752.

Tan, M. (1993). *Multi-agent reinforcement learning: Independent vs. cooperative agents.* In Proc. of the 10 international conference on machine learning, vol 337. Amherst, MA.

Outline

- so far we have focused on single agent DM
- multi-agent DM presents some challenges and complications
- we start from one-shot DM and later briefly mention sequential DM problems

Lecture's aim: briefly outline possible extension of DM to multi-agents within game theory framework.

..details see the course 01TEH Teorie her (T.Kroupa)

We have intelligent rational agent. What else?

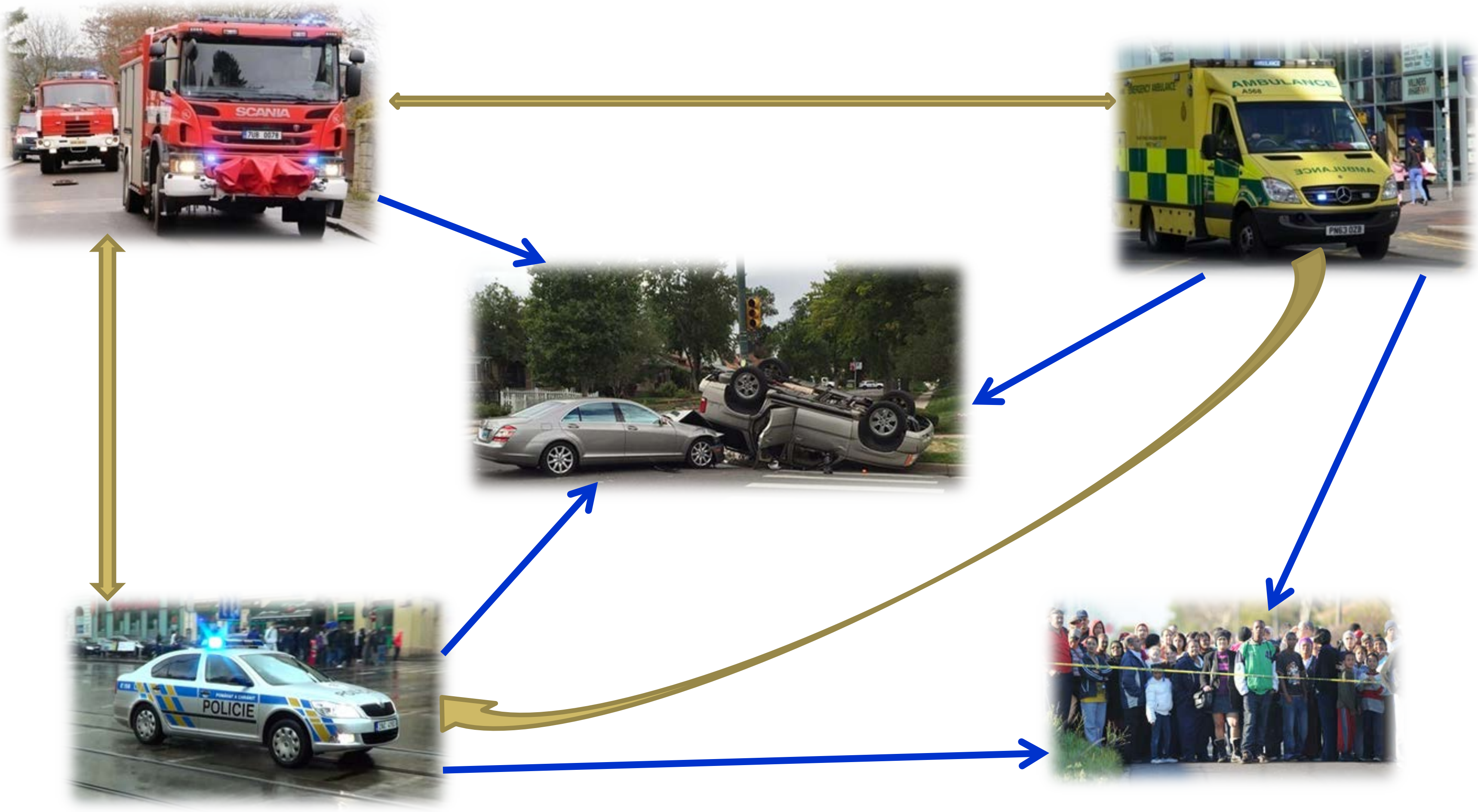
A single-agent DM has been considered so far. But

- Real life is multi-agents' *environment*: no one can act to achieve goals without respecting others actors.
- Some goals can only be achieved by interacting with others.
- (Computer) systems no longer stand alone, but are networked into large distributed systems (e.g. Internet, transportation networks and mobile sensing networks, power networks, synthetic biological networks,...)

Example: Hromadná nehoda na dálnici D1 (březen 2008)

- bylo zraněno 30 osob
- počet poškozených vozidel 231 (2 autobusy, 98 kamionů a 131 osobních aut).
- doprava byla zastavena na 11 hodin
- 150 profesionálních záchranářů
- v 30km koloně bylo zablokováno kolem 20 000 lidí a 2000 kamionů
- 30 odtahových vozidel, 2 záchranné a 3 policejní vrtulníky
- 40 hasičů ze 7 profesionálních jednotek

Example: Multi-agent DM



Example: goals and tasks

Global aim of ambulance teams, fire brigades and police forces:

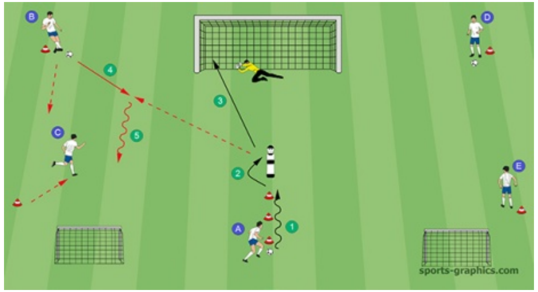
- Save as many lives as possible
- Save buildings and equipment whenever possible
- Manage disorder and panic

Local DM tasks of each team:

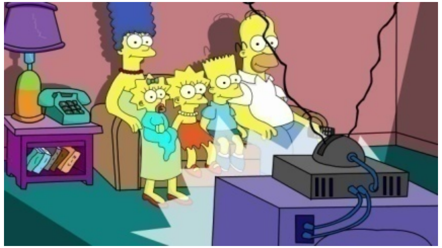
Ambulance team: how many ambulances from which hospital will go to which location? critical injuries- treat in-site; easier injuries - postpone or transport to hospitals.
Fire brigade: save vehicles/buildings or minimise fire spread
Police force: overall control the situation.

Complications:

- decisions are mostly *irreversible* and made under time pressure
- clear division of responsibilities, especially in overlapping space
- strong coordination of actions is crucial



MPDM issues



- Cooperation - working together as a team to achieve a *shared* goal.

  or when it is impossible to achieve the goal alone, or when cooperation bring better result (or sooner).
- Coordination - is *managing the interdependencies between activities.*
- Negotiation - *the ability to reach agreement* on issues of common interest. Typically involves offer and counter-offer and compromises made by agents.

Generally: ability to interact with other agents (and possibly humans) via cooperation, coordination, and negotiation is *social ability* of an agent.

Centralised or distributed?

- centralised: all social activities are coordinated by a central coordinator (agent with some privilegies)
- decentralised (distributed): agent is autonomously solves his social tasks while sharing knowledge with others. The ability to cooperate is intrinsic ability of the agent.
- mixed variants (depending on the class of tasks)

pros and cons

Knowledge sharing is about communication

- Neither of social activities can be solved *without knowledge sharing and communication*
- To communicate agents must have agreed a common set of terms - *ontology (a* formal specification of a set of terms)

- defining a common ontology determines the knowledge sharing => need a large effort (there exist special software tools)

the ability to communicate is not considered in the course.

Preferences (utilities) sharing

- is very important but much less developed topic. May efficiently help to better (implicit) coordination, cooperation and coalition formation.
- decrease conflicts and help to efficiently resolve them
- support coalition formation

MA DM: our tasks

1. to create an agent capable of independent, autonomous action in order to successfully carry out the tasks allocated
2. to create an agent capable of interacting (cooperating, coordinating, negotiating) with other agents in order to successfully carry out the tasks that we delegate to them
3. The same like 2 but when the other agents cannot be assumed to share the same interests/goals?

MA DM: our tasks

1. to create an agent capable of independent, autonomous action in order to successfully carry out the tasks allocated - done
2. to create an agent capable of interacting (cooperating, coordinating, negotiating) with other agents in order to successfully carry out the tasks that we delegate to them
3. The same like 2 but when the other agents cannot be assumed to share the same interests/goals?

Multi-agent system

A *multi-agent system* is a number of interacting agents

- are able to act in an environment
- which interact through communication
- have different goals and "spheres of influence" (may overlap)
- can be linked by other (technological or organisational) relationships.

In the most general case, agents act on behalf of (mostly human) users with *different* goals and motivations => mixed human-artificial

To successfully interact, they will require the ability to *cooperate, coordinate, and negotiate with each other* similarly as humans do.

Note: Every agent has own preferences (reward) and action spaces and optimises own reward.

Main tasks in multi-agent decision making

Multi-agent system: all agents influence environment $\Rightarrow$ the uncertainty is given by other agents too. And their decisions are influenced by ours.

- agent design

  What is a rational strategy given a game? (voting ,game, tax declaration, driving,..).

  Tool: game theory.

- mechanism design

  Given that agents behave rationally, what should the rules of the game be? (electing system; setting game rules; taxation system; designing driving laws;…)

  Tool: inverse game theory

Game Theory

Game theory is a formal tool for analysing interactions among a group of rational agents who behave strategically.

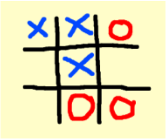
Group must have more than one decision maker (agent), otherwise it is a decision problem, not a game. For e.g solitaire is not a game.

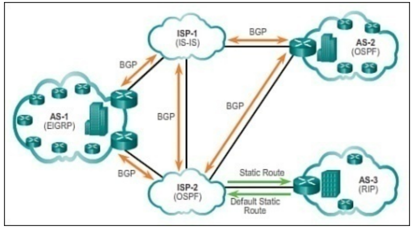
Game Theory (cont.)

Game theory is a formal tool for analysing *interactions* among a group of *rational* agents who behave *strategically*.

- interaction: each agent affects at least one other agent within the group
- rationality: an agent chooses its best action, i.e. optimising his objectives (utility)
- strategic: an agent respects how other agents influence the game.

Examples: games; auctions; trading; elections; any type of negotiation; driving in traffic; communication (routing protocols)





Main tasks in multi-agent decision making

Multi-agent system: all agents influence environment $\Rightarrow$ the uncertainty is given by other agents too. And their decisions are influenced by ours.

- agent design

  What is a rational strategy given a game? (voting ,game, tax declaration, driving,..).

  Tool: game theory.

- **mechanism design**

  Given that agents behave rationally, what should the rules of the game be? (electing system; setting game rules; taxation system; designing driving laws;…)

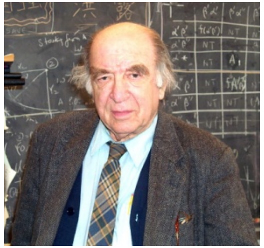
  Tool: inverse game theory

Mechanism design

‘’Inverse game theory’’ is a field of economics and game theory that defines the rules of the environment so that the collective welfare of all agents is maximised.

Conditioned that each agent adopts the game-theoretic solution maximising his own utility

Examples of related work in: *computer science* (network formation); *AI* (inverse RL), inverse *optimisation* (recovering the objective function from a solution); *economics* (revealed preferences of agent buying different goods), advertisement visiting and others



Leonid Hurwicz 1917 – 2008

Also known for *Hurwicz* criterion: in DM theory compromise between the maximax and maximin criterion, i.e. “optimistic vs pessimistic”balance.

Strategic (normal) game formulation

- a set of players (agents) $i=\{1,..,N\}$
- action set A_i for each player and join action set $A = X\ A_i$
- utility functions $u_i : A \rightarrow R$, expressing *agent's preferences* over the game outcomes, where $u_i\ (a)$ is utility of action a to player i

Table: e.g. matrix notation for 2players &2actions game

	C	D
A	u1/u2	u1/u2
B	u1/u2	u1/u2

Assumptions :

- players choose their actions independently (and simultaneously)
- players know the game structure, i.e. utilities and actions of all co-players

Note that assumptions are restrictive enough

Notations: actions $\equiv$ pure strategies; joint actions $\equiv$ pure strategy profiles;

utilities $\equiv$ payoffs; game's outcome $\equiv$ a numeric value for each player.

DM as a game

- Both agents influence the environment state
- each agent A has own strategy π_i^A that is dominating for A
- *Dominating* strategy provides the same or better result (reward) compare to any other strategy π_i^A for all $j=i$ and $\forall\, \pi_k^B$, $k=1,n$

Election example

- Agents: voting people (electors)
- Actions: possible votes for different candidates
- Outcome: set of all votes determines a winner (elected candidate)
- Utility function: describes preferences for each candidate

Wedding cars example

- Task: to deliver cars for bride (white) and groom (black)
- Agents: taxi services TSred and TSblue.
- Actions: {white car, Black car}
- What should TSs do:
  - both deliver white cars? agree?
  - deliver randomly?
  - p(black)=2/5, p(white)=3/5?

	Black car	White car
Black car	**0/0**	**10/12**
White car	**12/10**	**0/0**

Considerations how to solve the problem

Idea: use maximum expected utility. But *what will the co-player do*?

- agent’s own action is *not* enough to determine outcome/utility
- if agent had a distribution (beliefs) over the co-player’s actions, he could adjust his strategy.
- but his co-player can do *the same* => dependence of decisions is reciprocal!

Game theory provides models and solution concepts for general multi-agent interactions and a satisfactory solution has form of equilibrium, like (TSred gets white car; TSblue gets black car)

Dominance

- a strategy s for the agent *i* strongly *dominates* strategy s' if the outcome for s is better (for agent *i*) than the outcome for s', for every choice of strategies made by the other player(s)

  $\exists\ s : \forall s_{rest} \ \ u(s', s_{rest}) < u(s, s_{rest})$
- s is *dominating* and s' is *dominated* strategy
- A rational agent will never play a strictly dominated strategy!

Prisoner's Dilemma

If  () betrays a companion,  () will be free, while the companion gets 3 years in prison. If both betray each get 2years. If both cooperate and refuse to betray, each get 1year.

What should they do?

	Cooperate	Defect (betray A)
Cooperate	**1/1**	**3/0**
Defect (betray B)	**0/3**	**2/2**

years in prison

The rational and *dominating* decision is to defect.

BUT: the agent will be in a more difficult situation if co-player decides to cooperate.

Nash equilibrium and Pareto optimality

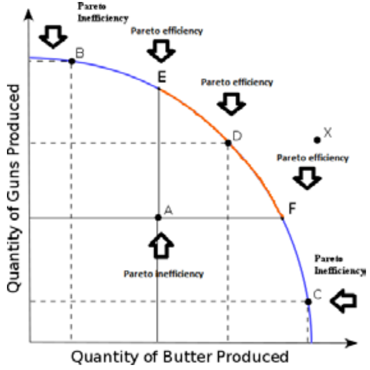
Nash equilibrium is solution that no player would benefit by changing own action while other players keep theirs unchanged.

John Forbes Nash Jr.
1928 – 2015

Pareto optimum (efficiency) there is no possibility to increase the player's payoff without decreasing payoff of at least one of his co-players. A state is Pareto dominated by some state if *all players* would prefer this state.

Vilfredo Federico Damaso Pareto
1848 –1923

Production possibilities frontier curve for producing "guns" and "butter“. A lies below the curve, denoting underutilized production capacity. B, C, D lie on the curve, denoting efficient utilization of production. X lies outside the curve, representing an impossible output for existing capital and/or technology.
(wikipedia.ru)



Prisoners Dilemma

A, B	U(.)	Pareto optimal	Nash equil.
(C,C)	1/1	yes	no
(C,D)	3/0	yes	no
(D,C)	0/3	yes	no
(D,D)	2/2	no	yes

Prisoner's Dilemma

If  () betrays a companion,  () will be free, while the companion gets 3 years in prison. If both betray each get 2years. If both cooperate and refuse to betray, each get 1year.

	Cooperate	Defect (betray A)
Cooperate	**1/1**	**3/0**
Defect (betray B)	**0/3**	**2/2**

years in prison

Why is PD a dilemma?

(defect, defect):

- is Nash equilibrium
- the *only* solution that is *not* Pareto optimal, but 'socially' optimal

(cooperate, cooperate) is a solution maximising common payoff (social welfare)

Common interest game (coordination game)

- A and B deciding whether to play tennis or chess

- each agent receives the same payoff
- no competition
- two pure Nash equilibria
- coordination is the main issue

How to solve the issue in practice?

	tennis	chess
tennis	1/1	0/0
chess	0/0	1/1

Common interest game (coordination game)

- A and B deciding whether to play tennis or chess

- each agent receives the same payoff
- no competition
- two pure Nash equilibria
- coordination is the main issue

How to solve the issue in practice?

What if it is **unbalanced** coordination game?

	tennis	chess
tennis	2/2	0/0
chess	0/0	1/1

Battle of the sexes (Luce and Raiffa,1957)

	opera	football
opera	2/1	0/0
football	0/0	1/2

- two-player coordination game
- non-cooperative
- two unfair Nash equlibria
- blue should select opera once he is sure red selects opera
- blue should select football once he is sure that red selects football
- .. and vice versa
- there is no dominating state

It is valid for non-sequential game otherwise the player making first step has an advantage (i.e. equilibrium plausible for him will be then selected)

Example: Matching Pennies Game

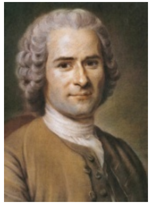
Players **A** *and* **B** *simultaneously choose the face of a* coin, either “heads” or “tails”. If they show the same face, then **A** *wins, while if they* show different faces, then **B** *wins.*

	Heads	Tails
Heads	**1/-1**	**-1/1**
Tails	**-1/1**	**1/-1**

What should the players do?

- if blue plays heads, red wants to play heads
- if red plays heads, blue wants to play tails
- if blue plays tails, read wants to play tails
- if read plays tails, blue wants to play heads…

What about random choice? 50-50? 30-70?



Jean-Jacques Rousseau
1712 – 1778

Stag hunt (Rousseau, 1775)

	Hunt 	Hunt 
Hunt 	**5/5**	**0/3**
Hunt 	**3/0**	**3/3**

Example of equilibrium selection:

If hunters cooperate ⇒  ; otherwise ⇒ 

Safer option: hunt alone for a hare.

There are two Nash equilibria (one is risk and another is payoff dominant):

(stag,stag), (hare,hare)

(stag,stag) ⇒ better payoff, but risky

(hare,hare) ⇒ lesser but guaranteed payoff

J.-J. Rousseau: Discourse on the Origin and Basis of Inequality Among Men, 1775

If the game is played repeatedly?

Repeated game :

- played repeatedly with the same players; different context are possible .
- the players may behave very differently than if the game is played just once (a one-shot game)
- players face the same task repeatedly, but each time with knowledge of the history all players' previous choices
- payoffs are additive over time
- strategies map history of play into (randomised) action choice

There are *finitely* (number of rounds is known and finite) and *infintely* (no predefined lengths, it terminates with some probability) *repeated games*

Equilibrium Selection

Complex topic which practical importance is proven in many applications (behavioural economics, social psychology, biology, recently in AI,..)

Choice of equiliblrium may be influenced by

- the *previous history* of the game (last time we were in opera, now lets go to football);
- social aspects or *convections/rules* (traffic in UK and CZ)
- *balance* in common interest (coordination) game
- *risk attitude* like in stag hunt game (cf. risk averse DM)
- *penalty* game
- *learning*

Pure & Mixed strategies

- pure strategy - deterministic selection (for one-shop it is a single action)
- mixed strategy - randomised choice (uncorrelated with other agents) that selects actions according to a given distribution over A . For 2 actions mixed strategy is $(p, a_1;(1-p),a_2)$

  Utility of the mixed strategy is $E(u(a)|p(a))$.

Two kinds of equilibrium:

- pure strategy equilibrium
- mixed strategy equilibrium

Equilibrium

- some games have no pure equilibria, only mixed (e.g. matching pennies)
- all finite games have some Nash equilibrium (mixed or pure)
- computation of Nash equilibria is generally very difficult except of special cases like two-players zero-sum games

Nash's theorem :

Any finite normal form game (finite number of players in which each player can choose from finitely many pure strategies) has at least one mixed strategy equilibrium.

Example: Matching Pennies Game

Players **A** *and* **B** *simultaneously choose the face of a* coin, either “heads” or “tails”. If they show the same face, then **A** *wins, while if they* show different faces, then **B** *wins.*

	Heads	Tails
Heads	**1/-1**	**-1/1**
Tails	**-1/1**	**1/-1**

- no pure equilibrium
- mixed equilibrium:

  if red plays Pr(H)=p; Pr(T)=1-p, then

  for $\forall p>0.5 \Rightarrow$ blue best response is Pr(T)=1;

  for $\forall p<0.5 \Rightarrow$ blue best response is Pr(H)=1.

  (0.5H,0.5T; 0.5H,0.5T) – the only Nash equilibrium. In other cases co-player will ‘’misuse’’ the situation and win.

Mixed equilibria: interpretations and critics

- allows for randomisation of actions to avoid exploitation (bluffing in poker, unfair penalty, etc..)
- represent conditions on stable beliefs about co-player style of playing (I believe that you believe that I believe that you play Pr(T)=0.5 ..)
- self-enforcing agreements, stable social convention
- learning: beliefs are based on past experience with co-playres, i.e. ability to predict co-player

- hard to find
- might not be unique
- human decisions may differ from the equilibria prediction (influence of fairness in bargaining game)

Perfect information games

- players know the past history (i.e. all previous moves and stochastic outcomes) cf. full observability MDP.
- similar to decision tree where each node is controlled by one player (corresponds complete game history); actions are edges and payoffs are at terminal nodes

- imperfect game $\neq$ incomplete info game (when full game structure is unknown, for instance one player does not know exact payoff of co-player)
- complete info = info about co-player, his actions, and payoffs

Markov games

- different “games” played at each round
- similar to repeated games
- at each time agents are in some state (as in an MDP)
- randomly selected actions determine:
  - immediate payoffs, and
  - (stochastic) state transition to the next round
- goal: to maximize total (or discounted) sum of payoffs

Iterated Prisoner's dilemma strategies

- A={cooperate, defect}
- s_t={cc,cd,dc,dd} – state of the game at time t
- $\pi : \otimes^{t}_{i=2} s_t \rightarrow A$
- Possible strategies for  :

$$\pi(s_t) = \begin{cases} c & if\ t = 1, \\ a_{t-1}^{blue} & if\ t > 1 \end{cases}$$

	Cooperate	Defect (betray A)
Cooperate	**1**/**1**	**3**/**0**
Defect (betray B)	**0**/**3**	**2**/**2**

Reinforce actions conditioned on game outcomes:

$$\pi(s_t) = \arg\min_{a} \mathrm{E}_T \left[\text{years in prison accumulated} \middle| s_t, a \right]$$

and update transition model.

Main tasks in multi-agent decision making

Multi-agent system: all agents influence environment $\Rightarrow$ the uncertainty is given by other agents too. And their decisions are influenced by ours.

- agent design

  What is a rational strategy given a game? (voting ,game, tax declaration, driving,..).

  Tool: game theory.

- **mechanism design**

  Given that agents behave rationally, what should the rules of the game be? (electing system; setting game rules; taxation system; designing driving laws;…)

  Tool: inverse game theory

Mechanism design: auctions example

Auction – mechanism for selling (allocation) some goods. Sotheby's ; Christie's; allocation of frequencies for mobile providers; processor operating; advertisement,..

Aim: to distribute resources/goods where they are needed most

Example: I've decided to give mark 'A' that of you, who needs it most.

How can I decide? What mechanism should I suggest to solve this?

Each of you has a specific utility value of 'A'. These values reflect your need. Each value is *your estimate* of the true value of 'A'.

Lets play a game: each of you write and seal your value (your move); the highest value wins 'A'. No one knows values of others => incomplete game.

Some types of auctions

- English auction (ascending bid) : starts from min; any bidder willing to buy proposes $b_t = b_{t-1} + b$, b is min rise; auction ends once no one willing to bid.

  dominating strategy: keep bidding until bid is <= than your value v; several bidders can cooperate and manipulate prices; strong bidder wins; high communication costs; high speed.
- Sealed-bid auction: bid is communicated to auctioneer; the highest wins.

  no simple dominant strategy; $b_{max}(v, \text{others}) + \varepsilon$, b_{max} prediction of maximal bid you can afford; dependence on predicted bids of co-players; more competitive then EA (player with highest v not necessary wins)
- Vickrey auction (sealed-bid second-price auction) . Like SBA, i.e highest bidder wins but the price paid is the second-highest bid ; dominating strategy is bid your value v. (used in Google's and Yahoo!'s online advertisement)

Mechanism design (cont.)

- Auctions illustrate other use of game theory, i.e. not searching for optimal decision, but *construction an optimal game* to ensure desirable outcomes.
- Applications:
  - election : find voting rules ensuring election result be acceptable for the majority
  - task allocation: how to allocate tasks ensuring fair load or efficient completion of the global task.
  - resource allocation: rules maximising group (society) benefit (taxes,m insurance)

Once preferences of the agents are unknown (or intentionally hidden) , use Bayesian game to solve.

Social choice

- *Social choice theory is concerned with group decision making. More general version of mechanism design problem.* Classic example of social choice theory: *voting.*
- given: a group of agents (society) having preferences over outcomes
- social choice function specifying proper outcome given the preferences of the population
- Formally, the issue is *combining preferences to derive* a *social outcome.*

Difference cf. mechanism design: preferences are qualitative orderings ; focus on scheme (like voting rules) satisfying some axioms; focus on computational aspects (communication complexity), manipulation; winner determination…

Applications: resource allocation in MAS, preference and rank learning in ML; ratings in advisory systems; voting; web-search data. Remember: big data are available!

Social choice: basic formulation

- Set of n alternative outcomes A={a1,a2,a3}
- m agents, each has preference ordering $\succ$ over A
- assume $\succ_k$ is a linear and strict order, k=1,..,m
- strategy (preference profile) $(\succ_1,\ldots,\succ_n)$ denote preference profile
- L is the set of linear orderings over A

- *social choice* function (SCF) SCF: $L^n \rightarrow A$ (i.e., consensus *winner*)
- *social welfare* function (SWF) CWF: $L^n \rightarrow L$ (i.e., consensus *ranking*)

(a1 $\succ$ a2 $\succ$ a3), (a3 $\succ$ a1 $\succ$ a2), (a2 $\succ$ a1 $\succ$ a3)

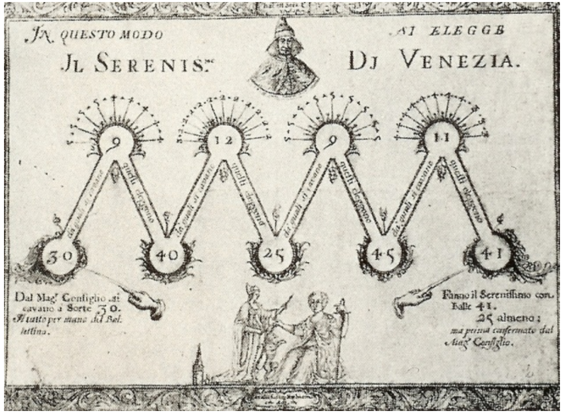
SWF: (a1 $\succ$ a2 $\succ$ a3); SCF: (a1)

Arrow thm

Elections of the Doge of Venice introduced in 1268

The complex electoral machinery aimed to minimise the influence of individual great families

- 30 members of the Great Council were chosen by lot
- The 30 are reduced by lot to 9
- The 9 elected 40 representatives
- The 40 were reduced by lot to 12
- The 12 representatives elected 25 representatives
- The 25 were reduced by lot to 9
- The 9 representatives elected 45 representatives
- The 45 were reduced by lot to 11
- The 11 elected 41 representatives
- These 41 actually elected the Doge of Venice.



details see wikipedia

Other aspects of social choice

- Behavioral social choice: design rules based on empirical preferences; modelling preference distributions (econometry, sociology)
- Preference aggregation: in complex (multi-issue) domains. how to combine them? exclude repetitive?
- Rank-based voting is complex and impractical (preference elicitation is important).
- Electing committees (multi-winner elections) : above ‘the line’ is not good especially in case of multiple preferences . All aspects should be considered.
- eliminating manipulation (adding/removing candidates); coalition; control of using mechanism, etc.
- numerical issues are important: how to approximate social choice function? decrease computation complexity?

Other applications of MAS

Agents as a tool for understanding human societies:

- Multi-agent systems provide a tool that may help shed some light on various kinds of social processes (via simulating artificial society).

Agents as a tool for understanding complexity

- SW engineering: for instance widely recognised that interaction is probably the most important single characteristic of complex SW.
- self-organising is an important property in physics/AI

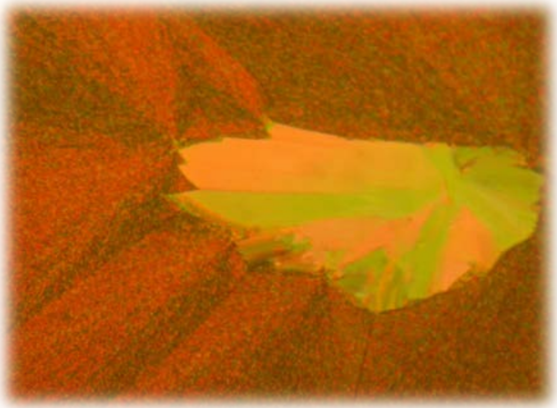
Multi-agent systems are of help where:

- control, data, expertise are distributed;
- centralised control is impossible or impractical; agents have competing/conflicting viewpoints or objectives.

Other interesting, but not covered issues of DM

- Self-organising and self-adaptive systems
- Rationality issues (Perfect and Bounded rationality)
- Social DM: modelling

Nature

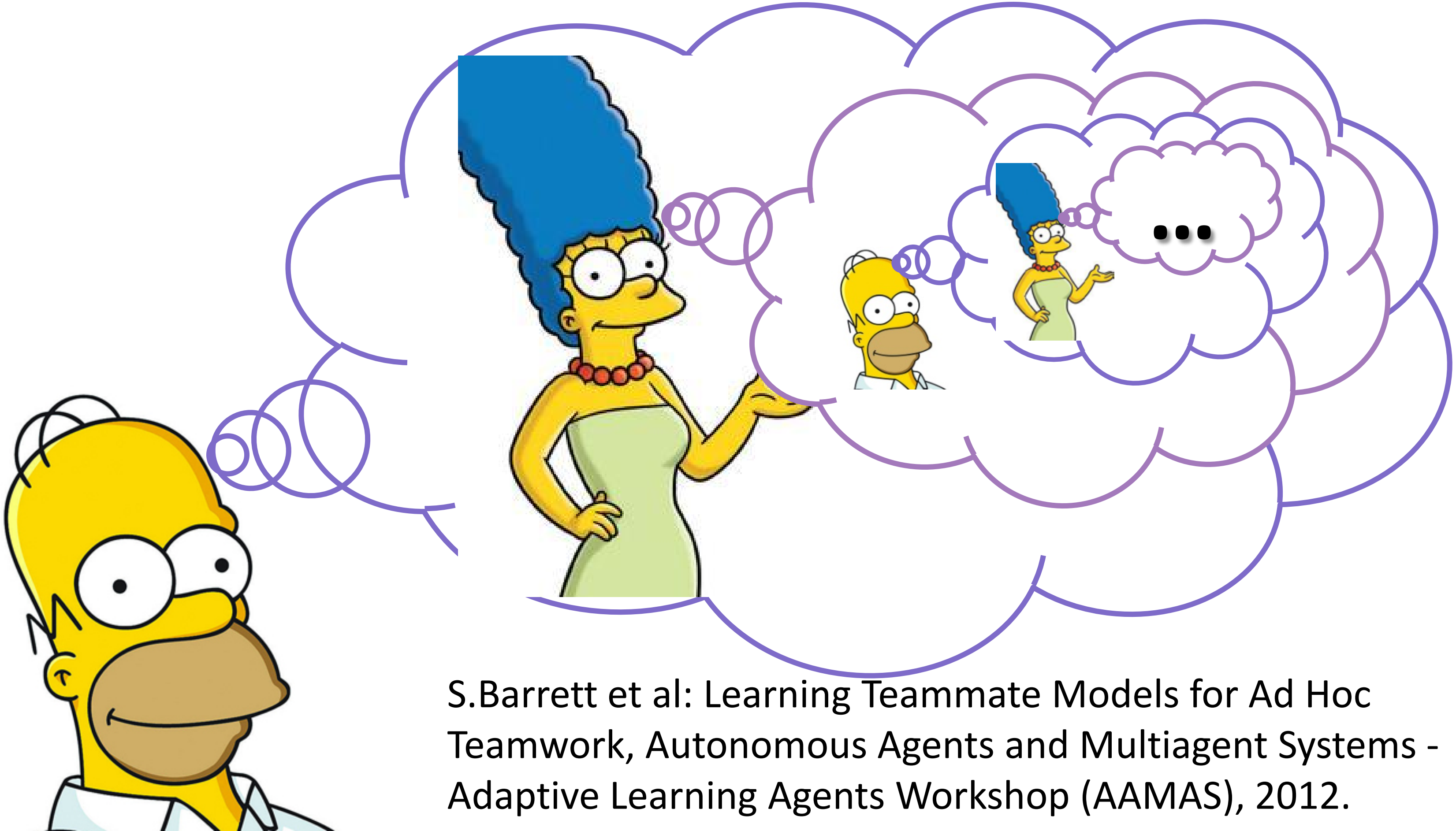


Emergence and Self-Organisation in Multi-Agent Systems

SO is a collection of interacting elements that gives rise to patterns of behaviors that individual elements are not capable of.

- Self-Organisation Mechanisms can be
  - Bio-inspired (flocking, reinforcement, cooperation, etc)
  - Social and Business/Economics Approaches (human collective behaviour, marketing,)
  - Fully Artificial Mechanisms (networking, self-configuring middleware, SW, etc)

Modelling of teammate and opponents



S.Barrett et al: Learning Teammate Models for Ad Hoc Teamwork, Autonomous Agents and Multiagent Systems - Adaptive Learning Agents Workshop (AAMAS), 2012.

Courses 01DRO1 and 01DROS wrap up

- Final exam info. Results.
- Other related courses
- AI research & jobs

Summary of Key Issues, see lecture 1

- Actions change state of the environment and enable other actions
- Sources of uncertainty:
  – knowledge of state of the world
  – effect of an action, exogenous inputs
  – behavior of other agents
- Action changes your knowledge i.e. provide some info though not *certainty* => value of information increases
- Effect of an action as well as preferences not known in advance
  – preference elicitation
  – learning (especially RL)
- Other agents in the environment
  – in cooperative settings: coordination of agents' activities
  – in competitive (fully, partially) settings: key is strategic/equilibrium effects

Topics covered so far:

- Review on probabilistic inference. Basics of sequential DM. Decision networks.
- Agent and its Environment. Rational DM. Formulation of DM task. Preferences ordering.
- Bayesian learning. Dynamic programming.
- Sequential DM: MDP and computation techniques for MDP. Value function (V) and action-value (Q-) function.
- Sequential DM: POMDP formalism.
- Learning for DM. Kalman Filter.
- Reinforcement learning. Model-free, model-based RL.
- Exploration vs. Exploitation

01DRO1 - Grading

- Midterm test (in class): weight 0.2
- Homework: weight 0.45
- Final exam: weight 0.35

  Note: each of (midterm test, homework, final exam) should be > *50%*!

- Bonus: full attendance (2 absences max) + *10%*

Grading: weighted average of the midterm test, homework and final exam.

Example: student Novak got marks:

midterm test – *60%;* homework – *70%;* final exam – *95%;* and bonus – *10%.*

Resulting mark: 0.2*60% + 0.45 *70% + 0.35*95% + 10%= 86.75%

"***B***" (dle klasifikační stupnice ČVUT)

Exam

- Weight 35%, in class
- Terms: will be announced on web-page and by email.
- **Material subject to examination:**
  - Slides for lectures 1-11
  - Mid-term test
  - Your course project (homework)
- No aids allowed, i.e. no notes, no lecture slides, no textbook, no PC/mobile, etc.
- Questions will focus on the *concepts and understanding*
- **Question types**: equations and definitions; explain some concept; numeric questions (e.g. calculate prediction using TD update); case study (mostly your course project) where you formalize something as an MDP
- Purpose is to help you see if you understand the major concepts covered in the course and able to use them

Other FJFI courses related to 01DRO1

- 01DRO2 – Dynamické rozhodování 2. Avanced continuation of 01DRO1.
- 01DROS – seminar Dynamické rozhodování
- 01STR – Statistická teorie rozhodování (Kůs)
- 01DYSY – Teorie dynamických systémů (Augustová)
- 01TEH – Teorie her (Kroupa)
- 01MAPR – Markovské procesy (Krbálek)
- 01UMIN – Pravděpodobnostní modely umělé inteligence (Vejnarová)

R&D: AI research and AI jobs

- Data Science: golden age of Machine Learning will end within a 3-5years
- Importance of decision making is gradually growing and will come into force within 5 years
- All large companies have strong AI R&D groups
  (Google, Microsoft, IBM, Amazon, Yahoo, Honewell, ..)
- Many small companies use AI (GoodAI, DataSentic,..) and develop some solutions
- 'Application-based' development: banks, insurance companies, state institutions, …

Lots to do for future research! Good luck!

PRÁÁÁÁZDNINY!!!