01DRO1
Dynamic Decision Making under Uncertainty

2017/2018

Course setup

- Tatiana V. Guy; guy@utia.cas.cz; ÚTIA 481; 266052254, 266052061.

- **Tuesdays: 15:30-17:30**, **T-212**
- 13 weeks; 2+0; **2** credits; **Exam**
- Required: AMSM; elective: MI, ASI

- consultations: Mondays, Fridays (email me in advance please), ÚTIA 481
- recommended literature and other course materials will be available on the course webpage http://staff.utia.cas.cz/guy/DRO12.html

01DRO1 - Course objectives

After this course you will be able to

- understand state-of-the-art of decision making (DM).
- master the probabilistic reasoning and optimization => broad applicability in AI, ML.
- read DM literature and literature in related fields (AI, ML, control)
- understand that any AI, ML or control task is indeed a *DM problem*
- formulate and formalise a DM problem
- look in a different way at many phenomena around you

01DRO1 - Grading

- Midterm test (in class)
- Homework
- Final exam

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

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

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)

Homework: **Possible options**

- Option A – Critical literature survey
- Option B –'Research' project
- Option C – Implementation of existing approach

Forms of output:

- report (6-8 pages), **or**
- 15' presentation with slides
  (should be presented by the author)

Homework: **Possible options**

- **Option A – Critical literature survey**
- Option B –'Research' project
- Option C – Implementation of existing approach

Forms of output:

- report (6-8 pages), **or**
- 15' presentation with slides
  (should be presented by the author)

Homework: **Critical literature survey**

What to do:

- Pick a problem (research topic) that *interests you*
- Search the literature for appropriate approaches to tackle this problem
- Survey and discuss the relative strengths of each approach (written)

Quality criteria:

to demonstrate a good understanding of the area and the ability to evaluate critically competing approaches;

ability to identify interesting future research directions/open problems.

You will have assistance with literature and topic selection

Recommended for those who wish to learn more about existing topic/method

Homework: **Possible options**

- Option A – Critical literature survey
- **Option B –'Research' project**
- Option C – Implementation of existing approach

Forms of output:

- report (6-8 pages), **or**
- 15' presentation with slides
  (should be presented by the author)

Homework: **'Research' project**

What to do:

- Choose a research problem (domain) that interests you
- Propose and investigate its solution (I will respect limited scope and time allocated)

Quality criteria:

to propose plausible/interesting solution; to show reasonable progress. Interesting (for you) topic can potentially serve for your Ing. thesis.

You will have assistance with topic and method selection

*Recommended for those who prefer to think over 'own' project and/or use this homework for future thesis (in that case I would recommend to discuss topic with **your supervisor** too).*

Homework: **Possible options**

- Option A – Critical literature survey
- Option B –'Research' project
- **Option C – Implementation of existing approach**

Forms of output:

- report (6-8 pages), **or**
- 15' presentation with slides
  (should be presented by the author)

Homework: **Implementation of existing approach**

What to do:

- Pick a problem that interests *you*
- Select an appropriate approach that solves your problem, implement and test it
- Discuss the strengths/weakness of the solution based on results

Quality criteria:

to select appropriate approach and construct a prototype implementation; show reasonable progress; discuss the strengths/weakness of the approach for your task; mention possible extensions and difficulties.

You will have assistance with problem and approach selection

Could be used for a particular task in your diploma thesis.

Recommended for those who incline to practical tasks and experimental work.

Homework: **logistic**

- Homework topic must be approved in advance. *You should submit your wish* (max 1/2 of A4!) by email due March 15, 2018.
- You will get my feedback by March 23, 2018. Earlier submission means earlier feedback/approval.
- Homework should be send by April 17, 2018. In case you did implementation, source files should be attached too.

Outputs: *either* report *or* presentation (should be presented by the author, length 15min).

Topics for inspiration, see course webpage and 01DROS. If you want to discuss a topic with me, email guy@utia.cas.cz and we'll arrange a meeting.

I strongly recommend you to start in advance.

New 01DROS:
Seminar Course on Dynamic Decision Making

- Elective course; **2+0z**; **2** credits.
- **Wednesdays , 15:30-17:30**, T-210
- 01DROS will support understanding how to design elements and methods inevitable for optimised decision making
- Practical problems are presented in collaboration with

and

New 01DROS:
Seminar Course on Dynamic Decision Making

How to develop DM methods theoretically and to apply them in practice.

- extends the topics learned in the lecture course 01DRO1
- describes actual topics and trends in DM, ML and AI
- introduces a number of practical tasks, for e.g.
  - practical introduction into ML and business intelligence
  - searching potential employees in web databases
  - taxi service
  - economic applications, futures trading
  - …

Highly recommended for everyone who has 01DRO1 as required course

New 01DROS:
Why should you attend the course?

You will ...

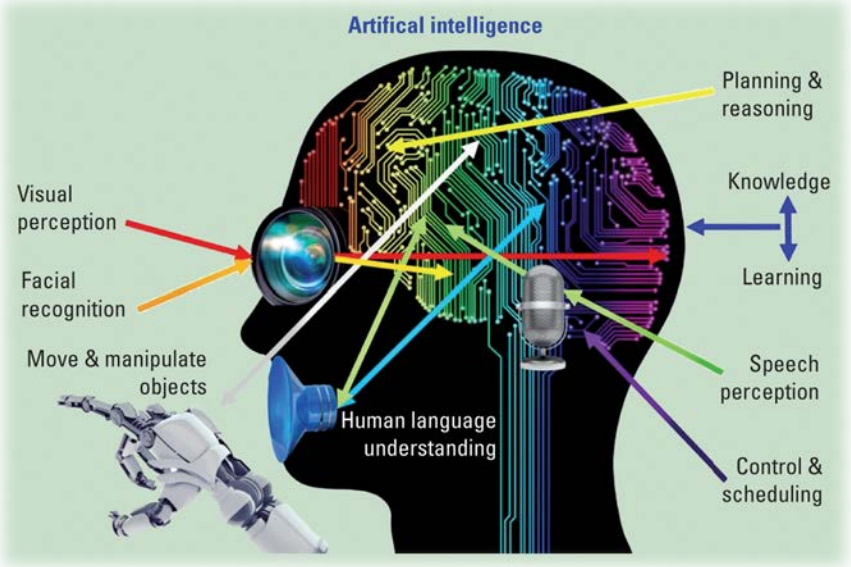
- get complementary information to 01DRO1
- get inspiration for your homework needed in 01DRO1
- discuss papers from the main DM, AI, ML conferences
- listen homework presentations of your colleagues
- listen guests lectures
- learn about real applications and solutions used
- …

Back to the today's course 01DRO1

Motivation

General need: Design, control and analysis of intelligent agents able to decide how to act in varying environments.

Sci-fi aim: creation of super intelligence.



Additional wish: (try) to mimic the behavior of biological organisms

Main ways

NATURAL (BIOLOGICAL) APPROACHES

Idea: copy and improve human being

- *Brain scan & Simulation*
- *Genetic Enhancement*
- *Brain Augmentation*

details on “brain in a vat” see movie *Matrix.*

ARTIFICIAL APPROACHES

Idea: design from first principles

- *Logic/language based*: expert/reasoning/cognitive systems.
- *Economics inspired*: utility, sequential decisions, game theory.
- *Cybernetics*: adaptive dynamic control.
- *Machine Learning*: reinforcement learning.
- *Information processing*: data compression ≈ intelligence.

DM goal

Each of artificial solutions is *too limited* but together they are powerful.

We still need theories that can guide our search for **intelligent rational** algorithms!

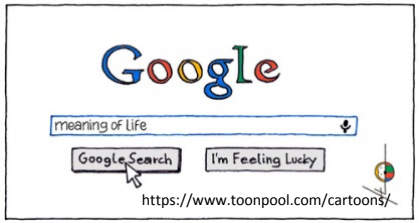
Decision making requires that the agent have knowledge (or beliefs) about its environment and its dynamics (incl. other agents), about its own abilities to observe and change the environment, and its goals and preferences.

01DRO1 will focus on probabilistic models and formalisation of DM under uncertainty incl. basic concepts of decision theory

Examples

- Search a solution
- Reasoning under uncertainty

- Natural Language processing



- Machine learning
- other

Relevant Research Fields

- computer science (artificial intelligence, machine learning),
- engineering (information theory, adaptive control),
- economics (rational agents, game theory),
- mathematics (statistics, probability),
- psychology (behaviorism, motivation),
- philosophy (reasoning, induction, knowledge).

Course 01DRO1 overview

Decision making under uncertainty:

- one-shot, sequential; single-agent, multi-agent
- probabilistic models of uncertainty, dynamic programming, Bayesian learning
- basic concepts of utility theory

Single-agent decision making:

- preferences, utilities: foundations, representations, elicitation
- knowledge elicitation; deliberation;
- sequential decision making, MDPs and POMDPs, Reinforcement learning

Multi-agent decision making (briefly, more detailed in 01DRO2 next term)

- coordination, cooperation, negotiation

Other concepts: risk attitude; certainty equivalence; framing effect

Lectures and readings with emphasis on perspective and discussion

This is a tentative schedule only. As the course progresses, the schedule will be adjusted.

Agent and Environment

Agent and Environment

Agent

(*knowledge*, *observations*) → *actions*

sensors

observations

actions

actuators

built-in knowledge, DM objectives

Environment

Agent: human; artificial device; both
Environment: part of the world

Physical vs. digital or virtual worlds make no difference

Environment

- Fully observable vs. partially observable
- Deterministic vs. stochastic
- Episodic vs. sequential
- Static vs. dynamic
- Discrete vs. continuous
- Single agent vs. multi-agent

Fully observable vs. partially observable

An agent's sensors give it access to the complete and correct state of the environment at each point in time.

The physical world (mostly) is partially-observable.

Deterministic vs. stochastic

Deterministic — the next state of the environment is completely determined by the current state and agent action. i.e. no uncertainty about the state.

Strategic – environment is deterministic except of actions of other agents.

- The physical world (mostly) is non-deterministic.
- Non-deterministic environments => harder to design an agent

Episodic vs. sequential

Episodic – decisions do not depend on previous decisions/actions.

Episodic environments - *simpler* as the agent can decide what action to perform based only on the current episode — it need not care about the interactions between this and future episodes.

Conversation is sequential.

https://www.execunet.com/

Static vs. dynamic

Static environment remains unchanged while an agent is deliberating except by the performance of agent actions.

The *physical world* = a highly dynamic environment.

Environment and Modelling

- Fully observable vs. partially observable
- Deterministic vs. stochastic
- Episodic vs. sequential
- Static vs. dynamic
- Discrete vs. continuous - number of actions/observations is fixed
- Single agent vs. multi-agent
- *Complete vs incomplete model – (learning)*

Hardest case: Partially observable, stochastic, sequential, dynamic, continuous, multi-agent environment ≡ Real world!

Environment: examples

Solitaire	Ultimatum Game	Bond's mission
Fully observable	Fully observable	Partially observable
Deterministic	Stochastic	Stochastic
Sequential	Sequential	Sequential
Static	Static	Dynamic
Discrete	Discrete	Continuous
Single-agent	Multi-agent	Multi-agent

Types of connection: Environment-Agent

Magic of Feedback or Closed-loop

Actuate
(gas pedal)

Measurement (sense) of car speed

Mr Bean
"computing"
control action

Magic of Feedback or Closed-loop

Idea of feedback:

- Compare the actual result with the desired result.
- Take actions based on the difference.

This simple but powerful idea is used everywhere: nature, science, engineering,...

- The opposite is feedforward or open-loop: make a plan and execute it.
- Feedback can be positive and negative.
- Feedback and feedforward are key ideas in control theory, more details see 01DYSY.

Feedback examples

Negative feedback

- Hungry cat => metabolism *slows down* to conserve energy and allow the cat to continue living with less food.
- Bicycle is *self-stabilising* if the velocity is larger than the critical velocity

Dreamstime.com

Positive feedback

- Markets with social influence: social influence => a Matthew effect i.e. “rich-get-richer” phenomena.
- Audio (acoustic) feedback may lead to endless instability

Main feedback properties

☺

- *robustness to uncertainty* and environment changes
- ability to *change dynamics* of the closed loop

☹

- feedback *instability* (e.g positive fb => the amplification on a microphone is turned up too high in a room)
- feedback *inherently couples* different parts of a closed loop (injects measurement noise into environment)
- feedback complicates => balance: costs vs. expected benefits

Closed-loop or Magic of Feedback

Feedback can “make good system from bad components”, i.e. it can

- make system insensitive to disturbances
- stabilise unstable system (bicycle)
- produce desired behaviour (e.g. non-linear components produce linear behaviour)

Feedback	Feedforward
Actions are computed based on the difference “actual output – desired output”	Make a plan of actions and execute it
Closed-loop control	Open-loop control
System-driven	“Pure” planning
Acts whenever there is a deviation of actual from desired => sensor noise is fed into system	Acts according to the planned sequence
Potential for instability	No risk of instability

Agent

(human being, controller, algorithm, device,..)

Intelligent Rational Agent

- Agent autonomously observes, decides and acts (close-loop interaction with environment)
- Rational agent acts to achieve best outcome when
  - changing the state of the environment (targeted influence)
  - discovering the state of the environment and make better decisions then (learn, describe)

- Agent should account for other agents in the environment

Intelligent Rational Agent

Rational agent acts to achieve best possible outcome

criteria of ‘best’ => performance measure for any realisations of behaviour and uncertainty

A rational agent chooses action that optimises the *expected value* of its performance measure given the actual data observed

Social ability

In multi-agent environments agent should have *social ability,* i.e. ability to interact with other agents via:

- *coordination* – influencing the interdependencies between actions of different agents (that can compete)
- *cooperation* is coordination of agent's actions to achieve a common (or shared) goal
- *negotiation* is reaching agreement on common issues (compromise)

https://www.wikihow.com/Score-A-Goal-In-Soccer

DM problem formalisation

needs description (from the agent's point of view) of:

- environment
- possible actions (actuators)
- possible observations (sensors)
- DM goal (performance measure -utility, loss, reward)

Note that **rationality** considers exploration, learning, adaptation

Prescriptive theory of DM under uncertainty

The theory helps agent

- to solve DM problems in dynamic, uncertain and randomly responding world
- to select strategy generating good actions when judged according personal preferences
- to avoid provably bad strategies
- to orient in myriads existing/possible ways of addressing DM
- to recognise “personally” conditioned sources of inevitable failures
- to have unifying framework for solutions of subtasks like learning, preference elicitation, etc.

Agent and Environment

Agent

(*knowledge*, *observations*) → *actions*

sensors

observations

Environment

actions

actuators

built-in knowledge, DM objectives

Agent: human; artificial device; both
Environment: part of the world

Physical vs. digital or virtual worlds make no difference

Key issues of Decision Making

Should Bond put the gun when soldiers held a gun to *his* head?

Good decisions depend on relative importance of *conflicting* or *competing* objectives => preferences are important

Example:

Bond ={*alive*, *injured*, *dead*}

mission = {*failed*, *successful*}. How to express preferences?

Problem: Not all objectives are easy to express by numbers!

Key issues of Decision Making

Should Bond respect soldiers' preferences?

Decisions should reflect Bond's *preferences* (generally: preferences of user on whose behalf artificial agent is acting)

artificial agent acts on behalf of human => no "preprogramming" (see bargaining/bidding agent; medical assistant; scheduler; etc.)

Preference elicitation/assessment needed if agent decides itself

Key issues of Decision Making

Bond's goal: take the bag and escape safely**.**

But: what if soldiers are watchful? what if gas balloons are empty? if the hidden gun misfires? if…

Bond must prepare an appropriate course of actions *depending on the current situation*:

- state can change exogenously *(*uncertainty*)*
- effects of Bond's actions can be uncertain (endogenous uncertainty)
- course of Bond's actions should be conditional (*policy, not plan) => closed loop!*

Key issues of Decision Making

Bond got new info (unknown to MI6) and should decide himself whether continue or interrupt his mission.

Decisions reflect tradeoffs between likelihood of outcomes and preferences over them.

Example: a student needs *2* credits and decides on selecting: 01DRO2 (*2*credits) or 01NEUR2 (*3*credits)? Prob(01DRO2==pass)=0.5;0.6? Prob(01NEUR2==pass)=0.3;0.4?

Consider that preparing for 01NUER2 is harder. How will your result change?

Key issues of Decision Making

Bond prefers to do his job perfectly. Executing his mission Bond caused an international accident and is officially grounded.

Decisions should respect *immediate* and *long-term consequences of actions* (as well as long-term objectives)

Example: do you feel the tasks below are the same?

Stop consuming alcohol if prob of health problems in 1 year is 0.2?

Stop consuming alcohol if prob of health problems in 20 years is 0.2?

Key issues of Decision Making

- Agents 007 and 009 have a shared mission.

  What are their actions?

- 007 is MI6 agent, while 009 has already defected from MI6

  What are their actions?

Decisions should reflect (anticipated) behavior of other agents

- coordination, cooperation, possible competition
- equilibrium, transferable utility
- elicitation and motivation

Summary of Key Issues

- 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

Conclusion

- Intelligence can have many faces ⇒ *formal definition difficult*
- Real world is terrible: partially observable, stochastic, sequential, dynamic, continuous, multi-agent, but luckily structured, ...
- Your generation should solve the problem of rational agents in uncertain worlds for the unknown environmental probability distribution.

No *Matrix* please!

Tasks to think:

Decide on the environment properties:

- crossword solving;
- photo analysis;
- medical diagnostics;
- car driving (auto-pilot);
- temperature controller in a room;
- satellite navigation system.

Example: a student needs *2* credits and decides on selecting: 01DRO2 (*2*credits) or 01NEUR2 (*3*credits)? Prob(01DRO2==pass)=0.5;0.6? Prob(01NEUR2==pass)=0.3;0.4? Consider that preparing for 01NUER2 is harder. How will your result change?

Tomorrow:

01DROS: Seminar Course on Dynamic Decision Making

- Elective course; **2+0z**; **2** credits.
- **Wednesdays, 15:30-17:30**, T-210
- Highly recommended for everyone who has 01DRO1 as a required course