# Decision making under uncertainty

## Course overview

Christos Dimitrakakis

October 29, 2013

## The problem of decision making under uncertainty

- Modelling our uncertainty about the world $\Rightarrow$ learning
- Optimising our decisions given our knowlege $\Rightarrow$ planning

## The problem of decision making under uncertainty

- Modelling our uncertainty about the world $\Rightarrow$ learning
- Optimising our decisions given our knowlege $\Rightarrow$ planning

## Applications and related problems

- Optimisation: robust decisions, efficient search, planning.
- AI: modelling, learning from interaction and/or demonstration.
- Economics: Mechanism design, behavioural modelling.
- Security: Cryptography, Biometrics, Intrusion detection and response
- Biology and Medicine: Automatic experiment design, clinical trials, congitive science.

## The problem of decision making under uncertainty

- Modelling our uncertainty about the world $\Rightarrow$ learning
- Optimising our decisions given our knowlege $\Rightarrow$ planning

## Applications and related problems

- Optimisation: robust decisions, efficient search, planning.
- AI: modelling, learning from interaction and/or demonstration.
- Economics: Mechanism design, behavioural modelling.
- Security: Cryptography, Biometrics, Intrusion detection and response
- Biology and Medicine: Automatic experiment design, clinical trials, congitive science.

Planning, learning and the exploration-exploitation trade-off

# The exploration-exploitation trade-off

## Example (Selecting a restaurant)

- You usually go to *Les Epinards*.
- You heard that *King's Arm* is really good!
- It's Friday. Do you:
  - Go to *Les Epinards*?

# The exploration-exploitation trade-off

## Example (Selecting a restaurant)

- You usually go to *Les Epinards*.
- You heard that *King's Arm* is really good!
- It's Friday. Do you:
  - Go to *Les Epinards*?
  - Call *King's Arm* to reserve?

# The exploration-exploitation trade-off

## Example (Selecting a restaurant)

- You usually go to *Les Epinards*.
- You heard that *King's Arm* is really good!
- It's Friday. Do you:
  - Go to *Les Epinards*?
  - Call *King's Arm* to reserve?
  - Check the menu of *King's Arm* and then decide?

# The exploration-exploitation trade-off

## Example (Selecting a restaurant)

- You usually go to *Les Epinards*.
- You heard that *King's Arm* is really good!
- It's Friday. Do you:
  - Go to *Les Epinards*?
  - Call *King's Arm* to reserve?
  - Check the menu of *King's Arm* and then decide?

## The exploration-exploitation trade-off

- Exploit knowledge about the world to gain a known reward.
- Explore the world to learn, potentially getting less or more reward.
- Arises when data collection is interactive.

## Formalising decision problems

- How do our decisions depend on what we want?
- How do we weigh alternatives?
- Is there a good concept of rationality?

## Formalising decision problems

- How do our decisions depend on what we want?
- How do we weigh alternatives?
- Is there a good concept of rationality?

## Beliefs, learning and planning

- How can we express belief and how does belief change?
- How might we make decisions according to our beliefs?
- What if our decisions can affect our beliefs?

## Formalising decision problems

- How do our decisions depend on what we want?
- How do we weigh alternatives?
- Is there a good concept of rationality?

## Beliefs, learning and planning

- How can we express belief and how does belief change?
- How might we make decisions according to our beliefs?
- What if our decisions can affect our beliefs?

## Why decision theory?

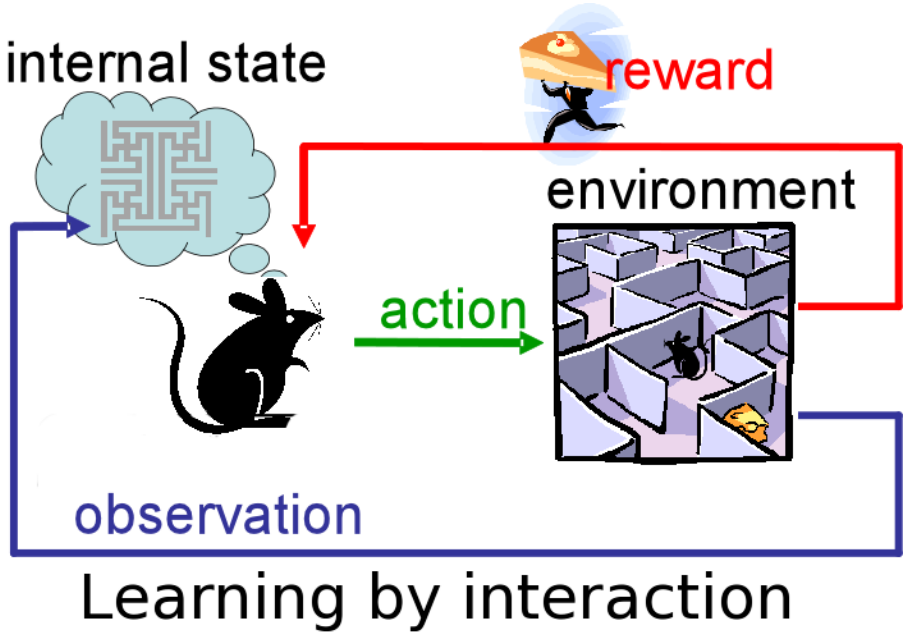
- Formalising trade-offs makes problems well-posed.
- Better overall solutions could be found.
- We may ignore non-essential aspects.

## The reinforcement learning problem

Learning to act in an unknown world, by interaction

### The interaction with the world

- The agent takes actions.
- The world generates observations.
- The agent receives rewards.




## Goal

Maximise total reward during the agent's lifetime:

- Fundamental problem in artificial intelligence.
- Connections to animal learning.
- Linked to experiment design, optimisation, game theory.

# Outline

* Probability refresher.
1. Subjective probability and utility.
2. Decision problems.
3. Estimation.
* Hypothesis testing.
4. Sequential sampling and optimal stopping.
5. Automatic experiment design and bandit problems.
6. Reinforcement learning I: Markov decision processes and fundamental algorithms.
7. Reinforcement learning II: Stochastic and approximation algorithms
8. Reinforcement learning III: Generalised problems.
9. Project meeting.
10. Reinforcement learning IV: Bayesian algorithms
11. Reinforcement learning V: Bandit algorithms and regret
12. Project meeting.
13. Learning with expert advice
14. Learning by demonstration; Preference Elicitation

# Assessment

## Exercises and feedback: 40%

- Exercises after every unit.
- Exercise sets include feedback form.
- Necessary for a good project!

# Assessment

## Exercises and feedback: 40%

- Exercises after every unit.
- Exercise sets include feedback form.
- Necessary for a good project!

## Participation: 10%

- Active participation in the course.
- Corrections on course notes.

# Assessment

## Exercises and feedback: 40%

- Exercises after every unit.
- Exercise sets include feedback form.
- Necessary for a good project!

## Participation: 10%

- Active participation in the course.
- Corrections on course notes.

## Project: 50%

Competition, presentation and report.

- Team competition using rl-glue socket API.
- Each team codes:
  - An environment (test-bed).
  - An agent.
- Agents are evaluated on all environments.

# Themes

- Models for representing belief and preferences.
- Algorithms for decision making.
- Fast optimisation.
- Applications in finance.
- Decision making in animals.
- Inferring preferences and beliefs.
- Automatic design of experiments.

[1] Dimitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[2] George Casella, Stephen Fienberg, and Ingram Olkin, editors. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, 1999.

[3] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning and Games*. Cambridge University press, Cambridge, UK, 2006.

[4] Morris H. DeGroot. *Optimal Statistical Decisions*. John Wiley & Sons, 1970.

[5] Marting L. Puterman. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, New Jersey, US, 1994.

[6] Leonard J. Savage. *The Foundations of Statistics*. Dover Publications, 1972.

[7] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.