facebook
Artificial Intelligence Research

Reinforcement Learning

Matteo Pirotta
Facebook AI Research

Acknowledgment

Extended version of the RL class **Alessandro Lazaric** (FAIR) and I are teaching in Paris

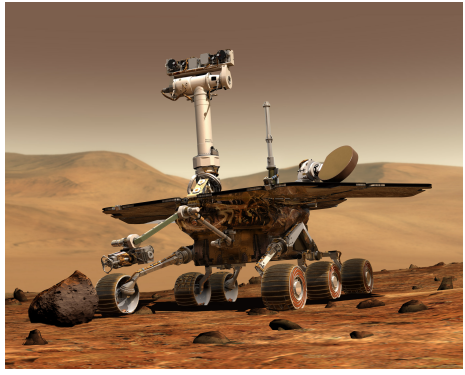
Special thanks to Alessandro for providing the initial version of the slides.

Teaching Assistants:
- Florian Strub (Deepmind)
- Omar Darwiche Domingues (INRIA Lille)

Why: Important Problems

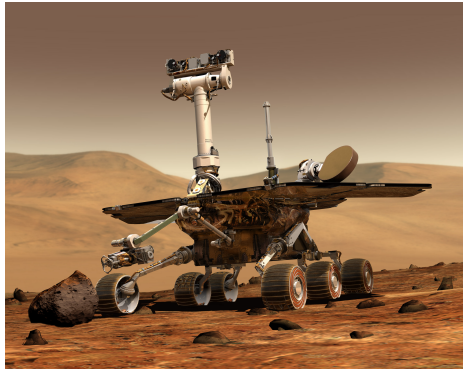
Why: Important Problems

- Autonomous robotics



Why: Important Problems

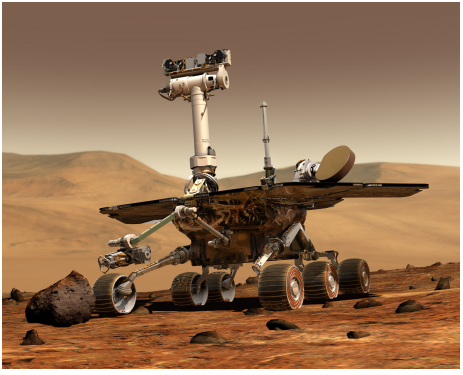
- Autonomous robotics



- *Elder care*

Why: Important Problems

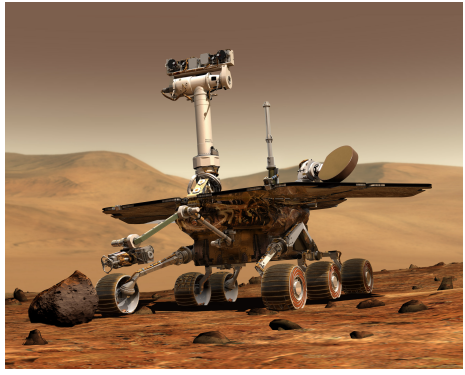
- Autonomous robotics



- *Elder care*
- *Exploration of dangerous environments*

Why: Important Problems

- Autonomous robotics



- *Elder care*
- *Exploration of dangerous environments*
- *Robotics for entertainment*

Why: Important Problems

- Robotics
- Finance

Why: Important Problems

- Robotics
- Finance

- *Trading execution algorithms*

Why: Important Problems

- Robotics
- Finance

- *Trading execution algorithms*
- *Portfolio management*

Why: Important Problems

- Robotics
- Finance

- *Trading execution algorithms*
- *Portfolio management*
- *Option pricing*

Why: Important Problems

- Robotics
- Finance
- Resource management

Why: Important Problems

- Robotics
- Finance
- Resource management

- *Energy grid integration*

Why: Important Problems

- Robotics
- Finance
- Resource management

- *Energy grid integration*
- *Energy market regulation*

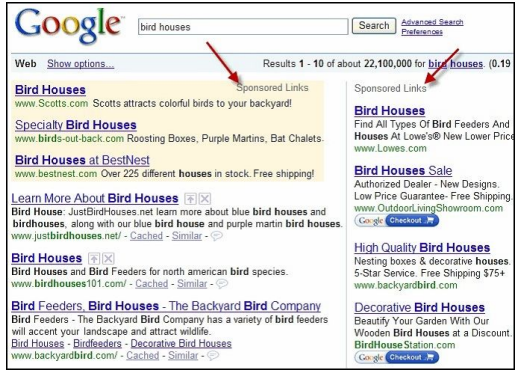
Why: Important Problems

- Robotics
- Finance
- Resource management

- *Energy grid integration*
- *Energy market regulation*
- *Energy production management*

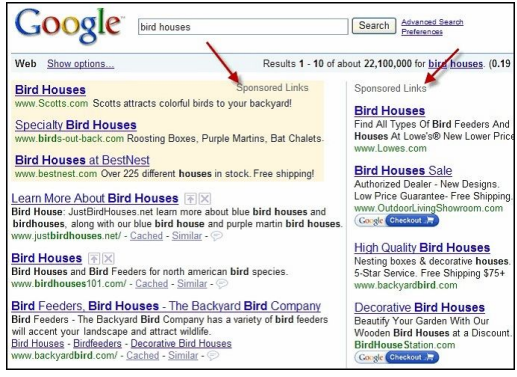
Why: Important Problems

- Robotics
- Finance
- Resource management
- Recommender systems



Why: Important Problems

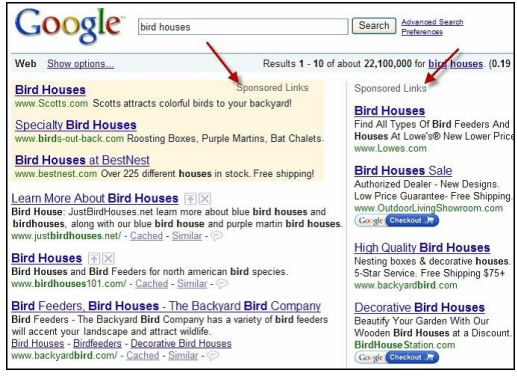
- Robotics
- Finance
- Resource management
- Recommender systems



- *Web advertising*

Why: Important Problems

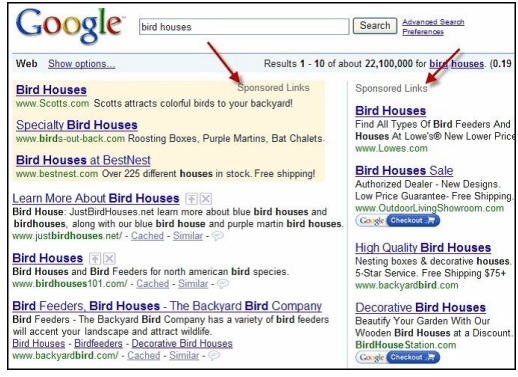
- Robotics
- Finance
- Resource management
- Recommender systems



- *Web advertising*
- *Product recommendation*

Why: Important Problems

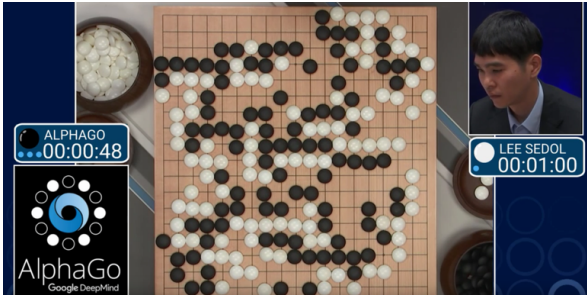
- Robotics
- Finance
- Resource management
- Recommender systems



- *Web advertising*
- *Product recommendation*
- *MOOCs / ITS*

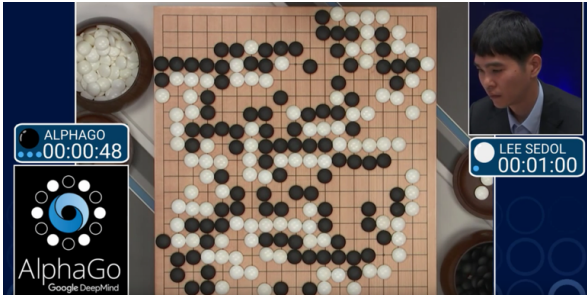
Why: Important Problems

- Robotics
- Finance
- Resource management
- Recommender systems
- Games



Why: Important Problems

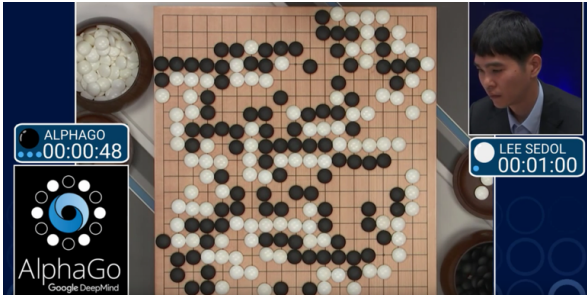
- Robotics
- Finance
- Resource management
- Recommender systems
- Games



- *Board games*

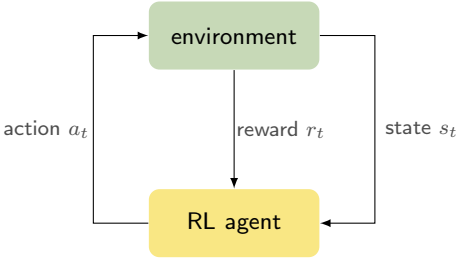
Why: Important Problems

- Robotics
- Finance
- Resource management
- Recommender systems
- Games



- *Board games*
- *Computer games*

What: Reinforcement Learning



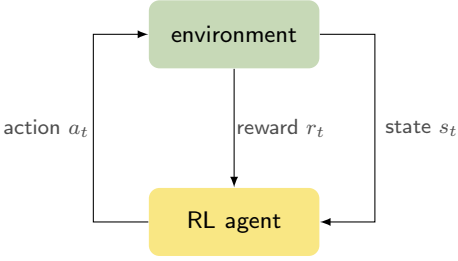
"**Reinforcement learning** is learning how to map states to actions so as to **maximize** a numerical **reward** signal in an unknown and **uncertain** environment.

In the most interesting and challenging cases, **actions** affect not only the immediate reward but also the **next situation** and all subsequent rewards (**delayed reward**).

The agent is not told which actions to take but it must discover which actions yield the most reward by trying them (**trial-and-error**)."

— Sutton and Barto [1998]

What: Reinforcement Learning



A framework for learning by interaction under uncertainty

Learn by sequentially interacting with the environment. The agent takes an action:

- receives an *instanteneous* reward
- the environment evolves to a new state

keywords:

Reward: feedback about the utility of an action

Value: cumulative reward the agent can get in a state

Policy: rule to select an action in a state

What: the Highlights of the Course

How to *model* an RL problem

- *What:* Markov decision process
- *Tools:* probability, processes, Markov chain

What: the Highlights of the Course

How to *model* an RL problem

How to solve *exactly* an RL problem

- *What:* Dynamic programming
- *Tools:* fixed point operators

What: the Highlights of the Course

How to *model* an RL problem

How to solve *exactly* an RL problem

How to solve *incrementally* an RL problem

- *What:* temporal difference, Q-learning
- *Tools:* stochastic approximation

What: the Highlights of the Course

How to *model* an RL problem

How to solve *exactly* an RL problem

How to solve *incrementally* an RL problem

How to solve *approximately* an RL problem

- *What:* approximate RL (policy gradient, TD-based methods, deep RL)
- *Tools:* supervised learning, optimization, deep learning

What: the Highlights of the Course

How to *model* an RL problem

How to solve *exactly* an RL problem

How to solve *incrementally* an RL problem

How to solve *approximately* an RL problem

How to *efficiently* explore in an RL problem

- *What:* multi-armed bandit problem
- *Tools:* concentration inequalities

What: the Highlights of the Course

How to *model* an RL problem

How to solve *exactly* an RL problem

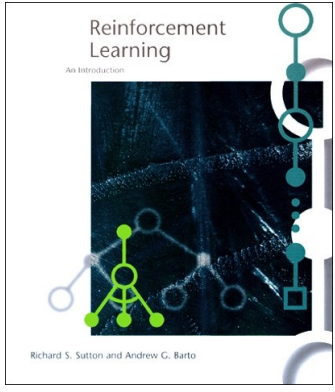
How to solve *incrementally* an RL problem

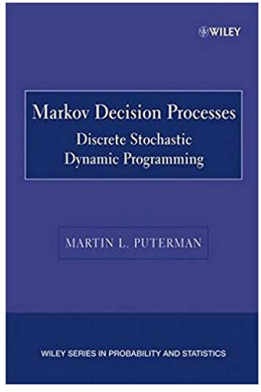
How to solve *approximately* an RL problem

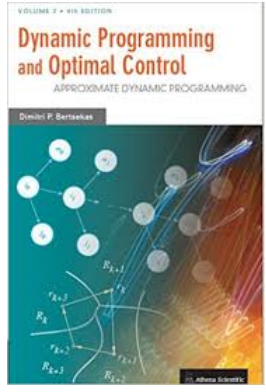
How to *efficiently* explore in an RL problem

With (simple!) examples from *resource optimization*, *trade execution*, *computer games*, *recommendation systems*.

Classical Books







Sutton and Barto [1998], Puterman [1994], Bertsekas [2007]

What: Machine Learning

RL is view as a subfield of machine learning

- Supervised Learning
  - Learn to make predictions based on samples containing *targets*
- Unsupervised Learning
  - Learn some latent structure in data
- Reinforcement Learning
  - Solve sequential decision problem (influence sample collection process)

What

1. Markov Decision Processes
2. Dynamic Programming
3. Reinforcement Learning: Tabular case
4. Approximate Reinforcement Learning: value-based
5. Approximate Reinforcement Learning: policy-based
6. Exploration: Bandits
7. Exploration: Reinforcement Learning

Friday (21th) 9:00-10:00 (am): Quiz
Friday (28th) 9:00-10:00 (am): Quiz

Thank you!

facebook
Artificial Intelligence Research

Bibliography

Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. II*. Athena Scientific, 3rd edition, 2007.

M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, Etats-Unis, 1994.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Bradford Book, MIT Press, Cambridge, MA, 1998.