A Quick Look at the "Reinforcement Learning" course

A. LAZARIC (*SequeL Team @INRIA-Lille*)
Ecole Centrale - Option DAD

Why

Why: Important Problems

Why: Important Problems

- Autonomous robotics

Why: Important Problems

- Autonomous robotics

- *Elder care*

Why: Important Problems

- Autonomous robotics

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

Why: Important Problems

- Autonomous robotics

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

Why: Important Problems

- Autonomous robotics
- Financial applications

Why: Important Problems

- Autonomous robotics
- Financial applications

- *Trading execution algorithms*

Why: Important Problems

- Autonomous robotics
- Financial applications

- *Trading execution algorithms*
- *Portfolio management*

Why: Important Problems

- Autonomous robotics
- Financial applications

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

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management

- *Energy grid integration*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management

- *Energy grid integration*
- *Maintenance scheduling*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management

- *Energy grid integration*
- *Maintenance scheduling*
- *Energy market regulation*

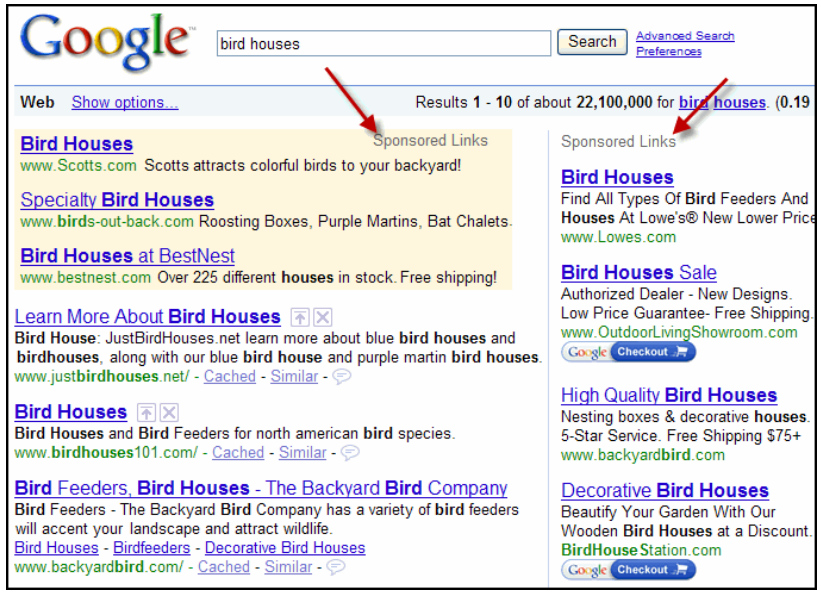
Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management

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

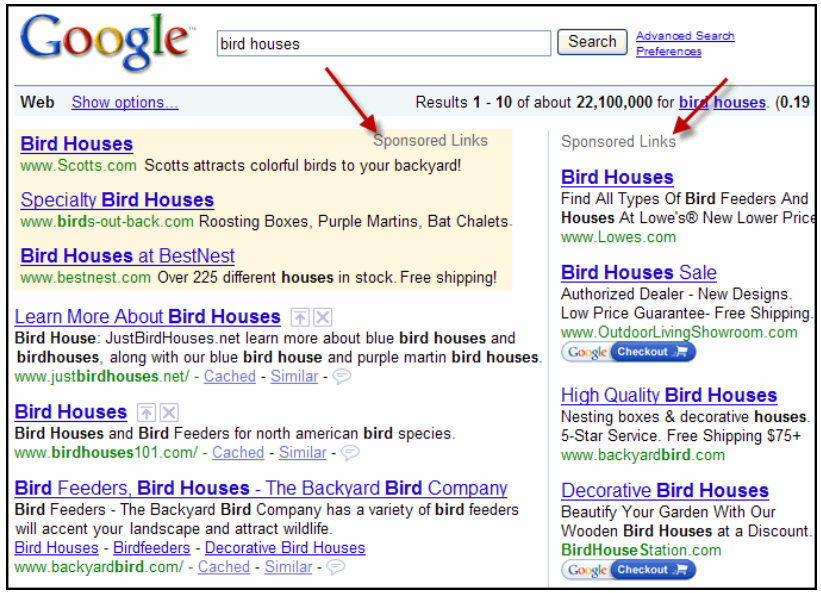
Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems



Why: Important Problems

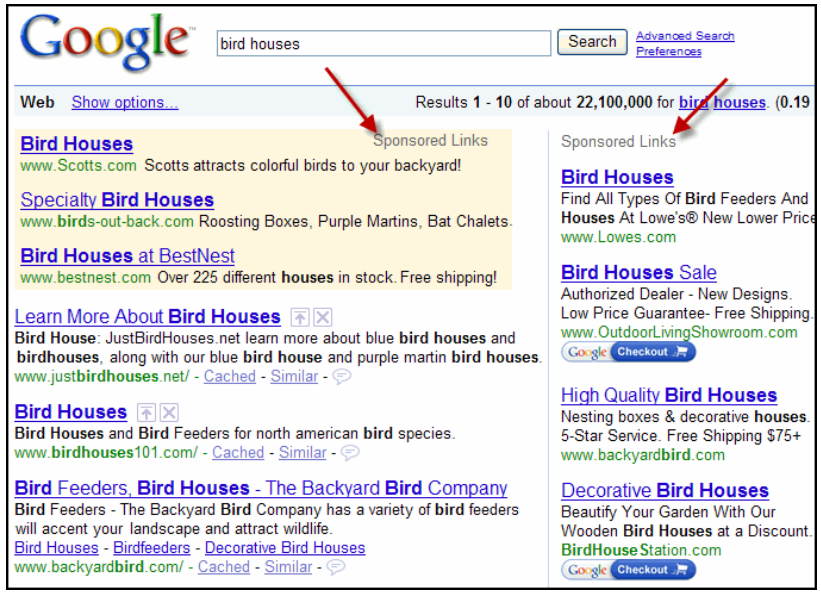
- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems



- *Web advertising*

Why: Important Problems

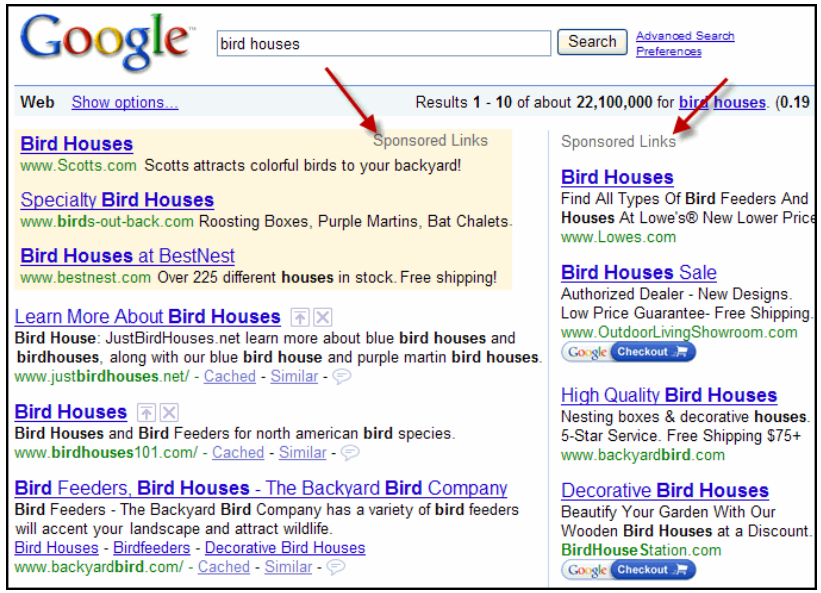
- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems



- *Web advertising*
- *Product recommendation*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems



- *Web advertising*
- *Product recommendation*
- *Date matching*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems
- Social applications

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems
- Social applications

- *Bike sharing optimization*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems
- Social applications

- *Bike sharing optimization*
- *Election campaign*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems
- Social applications

- *Bike sharing optimization*
- *Election campaign*
- *ER service optimization*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems
- Social applications

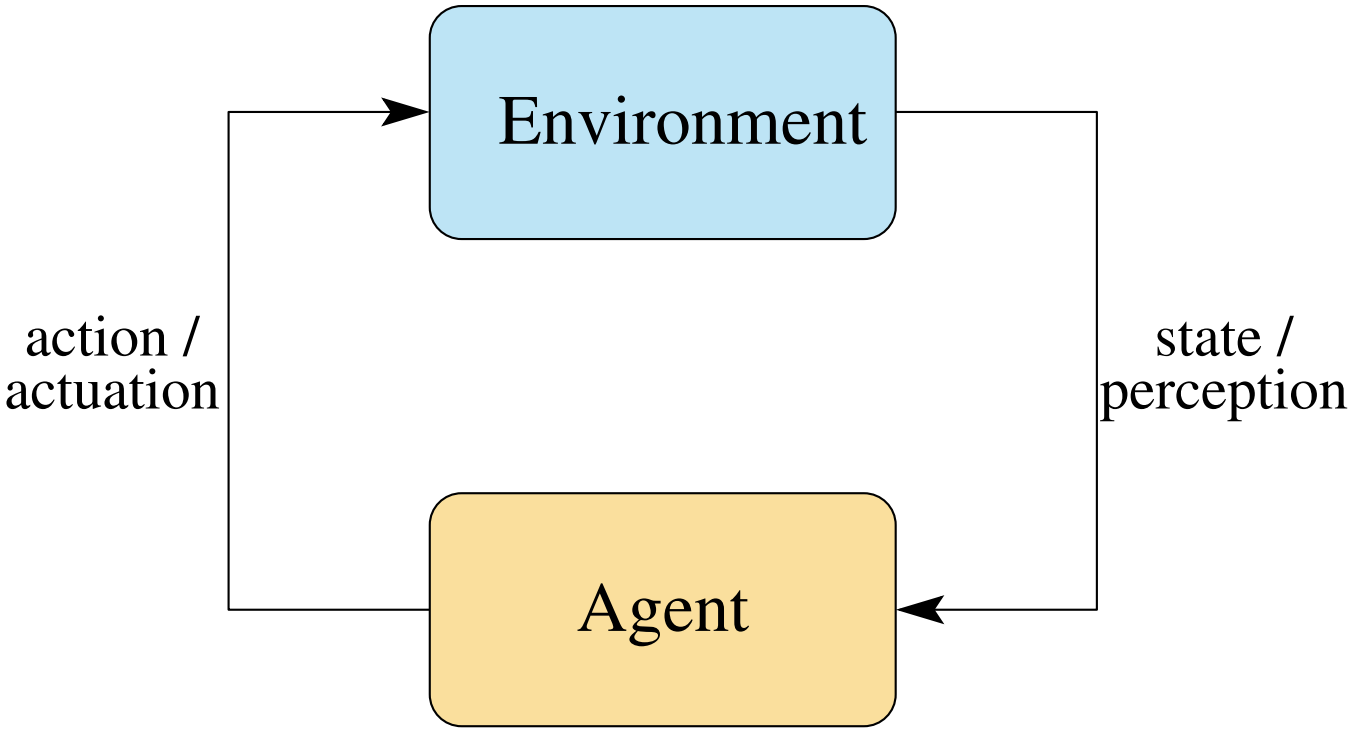
- *Bike sharing optimization*
- *Election campaign*
- *ER service optimization*
- *Resource distribution optimization*

Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems
- Social applications
- And many more...

What

What: Decision-Making under Uncertainty



How: Reinforcement Learning

Reinforcement learning *is learning what to do – how to map situations to actions – so as to maximize a numerical reward signal. The learner is not told which actions to take, as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them (**trial–and–error**). In the most interesting and challenging cases, actions may affect not only the immediate reward but also the next situation and, through that, all subsequent rewards (**delayed reward**).*

"An introduction to reinforcement learning",
Sutton and Barto (1998).

How: the Course

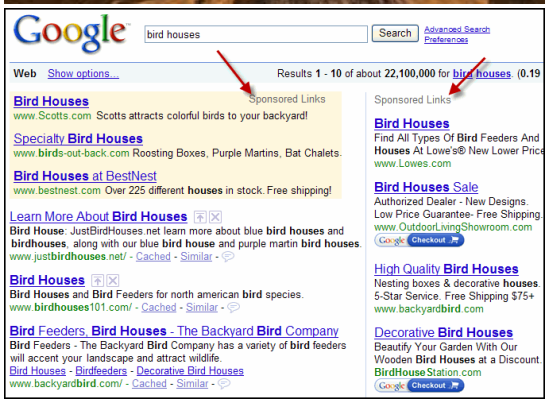
How: the Course

How: the Course

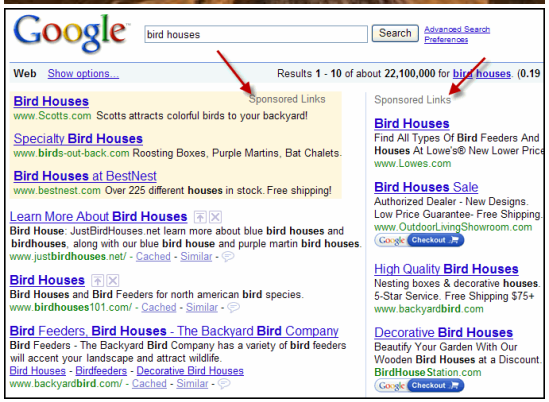
How: the Course



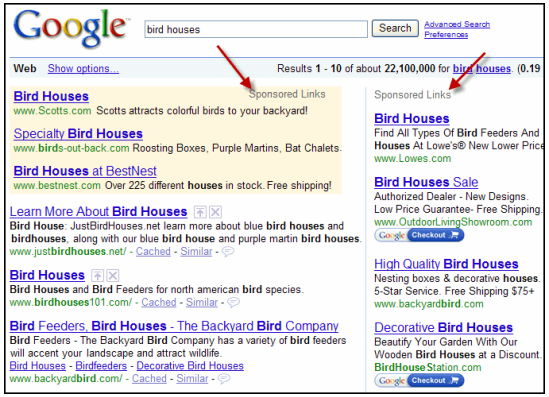
How: the Course

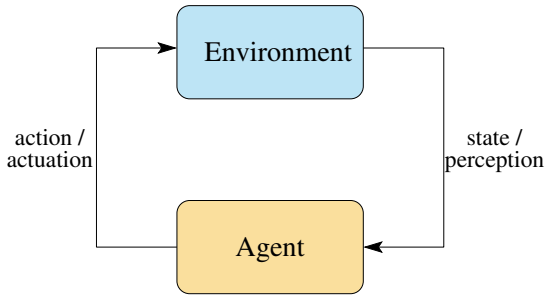


How: the Course

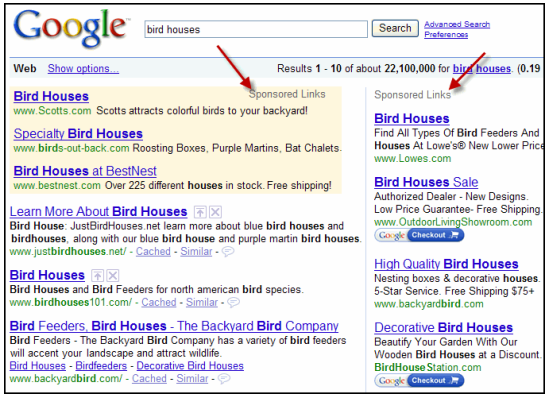


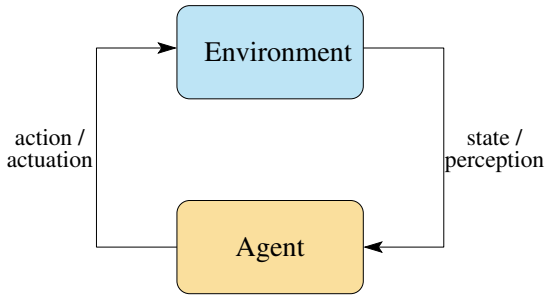
How: the Course





How: the Course





Formal and *rigorous* approach to
the RL's way to decision-making under uncertainty

What: the Highlights of the Course

How do we formalize the agent-environment interaction?

What: the Highlights of the Course

How do we formalize the agent-environment interaction?

How do we solve an MDP?

What: the Highlights of the Course

How do we formalize the agent-environment interaction?

How do we solve an MDP?

How do we solve an MDP "online"?

What: the Highlights of the Course

How do we formalize the agent-environment interaction?

How do we solve an MDP?

How do we solve an MDP "online"?

How do we effectively trade-off exploration and exploitation?

What: the Highlights of the Course

How do we formalize the agent-environment interaction?

How do we solve an MDP?

How do we solve an MDP "online"?

How do we effectively trade-off exploration and exploitation?

How do we solve a "huge" MDP?

Who

Lectures and Practical Sessions

Alessandro LAZARIC

SequeL Team
INRIA-Lille Nord Europe

alessandro.lazaric@inria.fr
researchers.lille.inria.fr/~lazaric/

When/What/Where

See planning on the website.

Evaluation

- Three homework (dynamic programming, multi-armed bandit, approximate dynamic programming): 2.5 points each.
- Review of literature with oral presentation: 12.5 points.

Reinforcement Learning

Alessandro Lazaric

alessandro.lazaric@inria.fr

sequel.lille.inria.fr