A Quick Look at the "Reinforcement Learning" course

A. LAZARIC (*SequeL Team @INRIA-Lille*)
ENS Cachan - Master 2 MVA

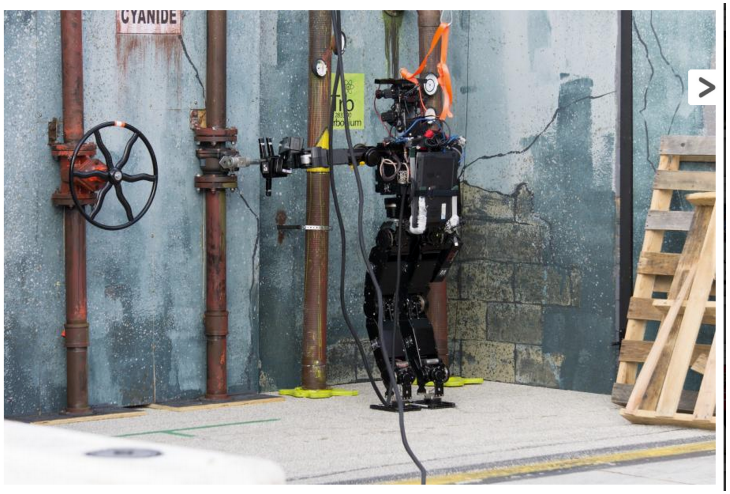
SequeL – INRIA Lille

Why

Why: Important Problems

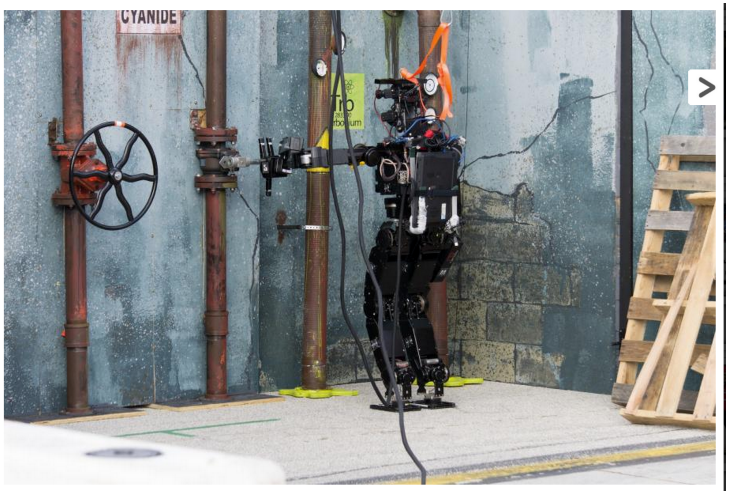
Why: Important Problems

- Autonomous robotics



Why: Important Problems

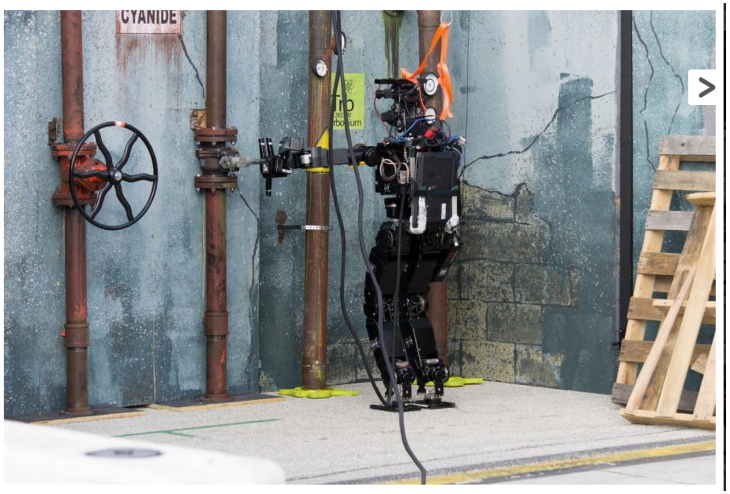
- Autonomous robotics



- *Elder care*

Why: Important Problems

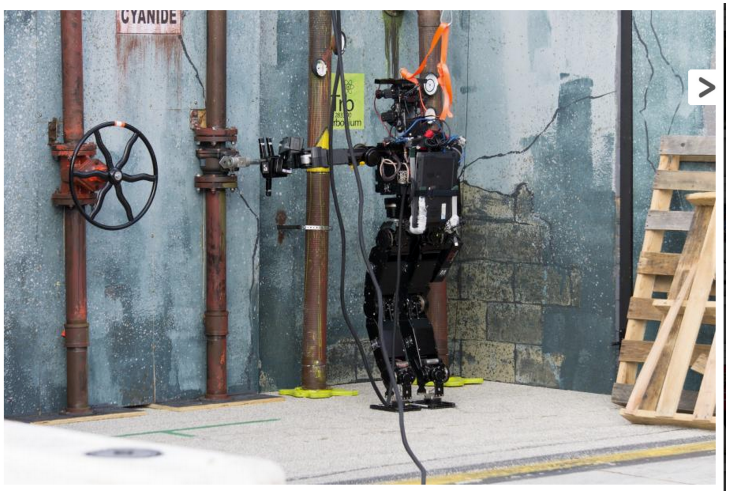
- Autonomous robotics



- *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*

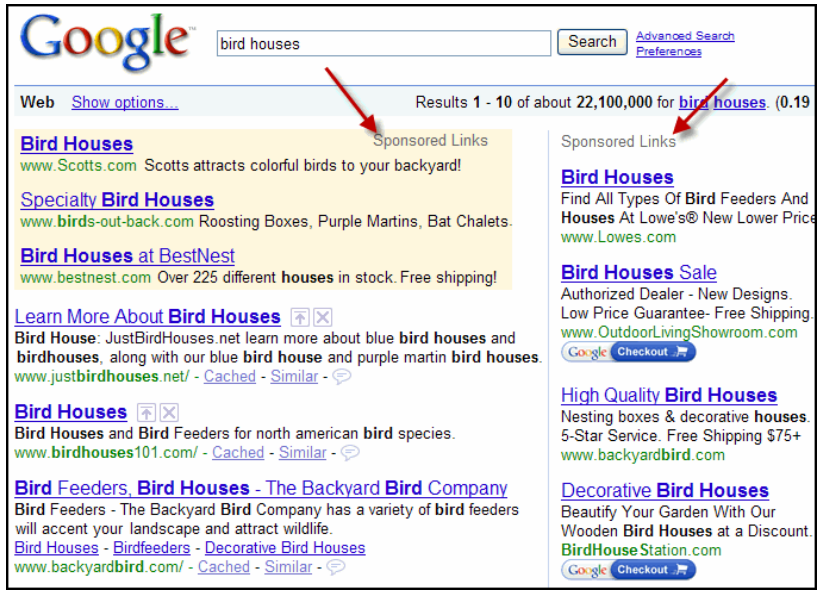
Why: Important Problems

- Autonomous robotics
- Financial applications
- Energy management

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

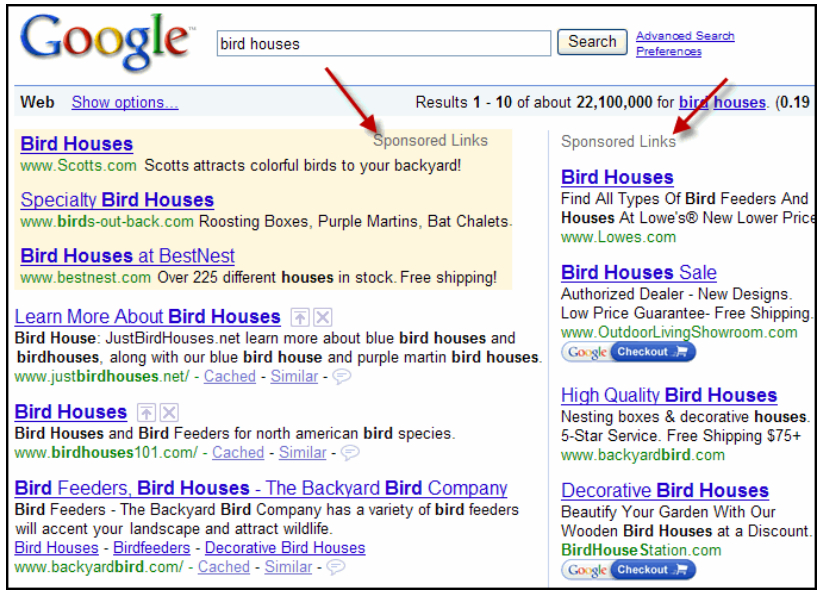
Why: Important Problems

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



Why: Important Problems

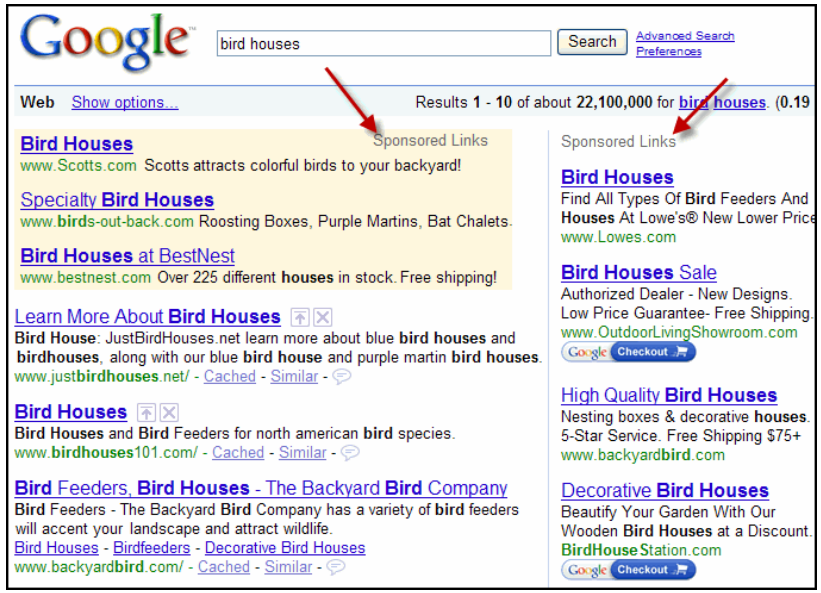
- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems



- *Web advertising*

Why: Important Problems

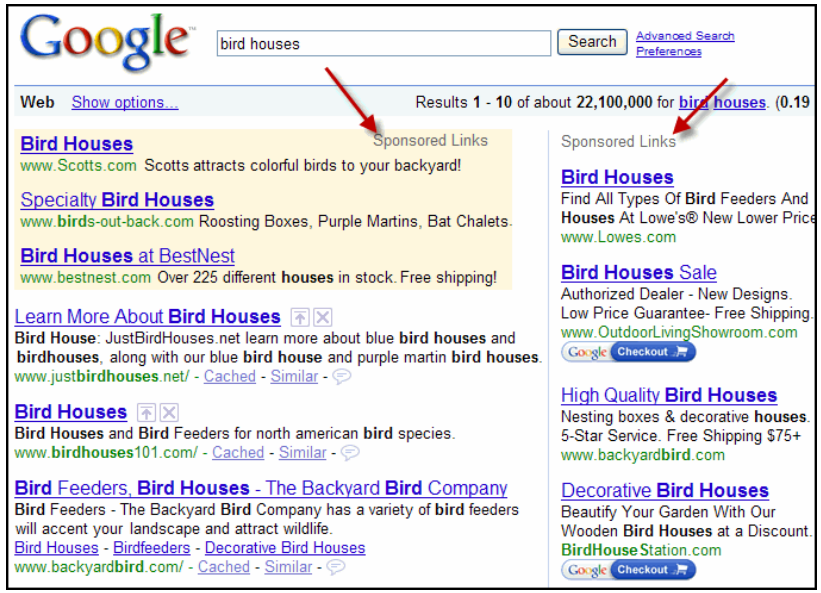
- Autonomous robotics
- Financial applications
- Energy management
- Recommender systems



- *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*

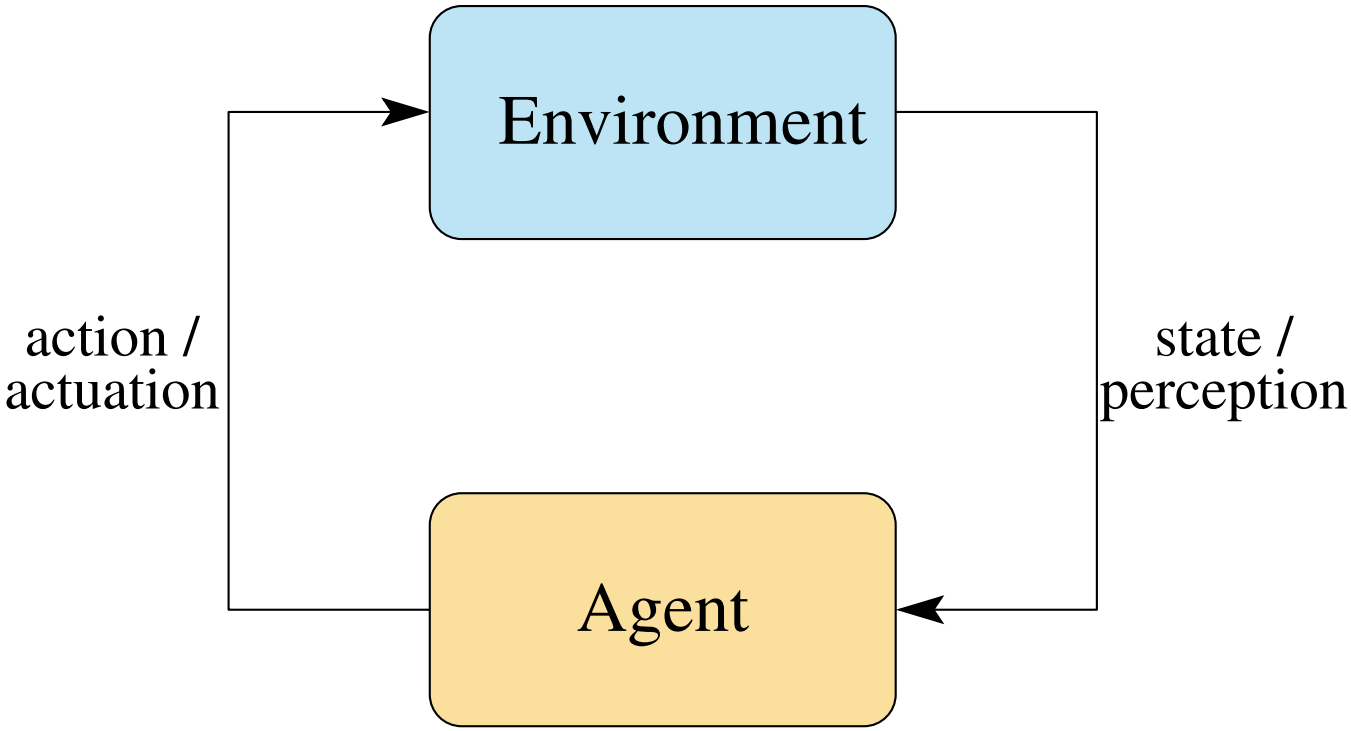
Why: Important Problems

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

- *Bike sharing optimization*
- *Election campaign*
- *ER service optimization*
- *Intelligent Tutoring Systems*

What

What: Decision-Making under Uncertainty



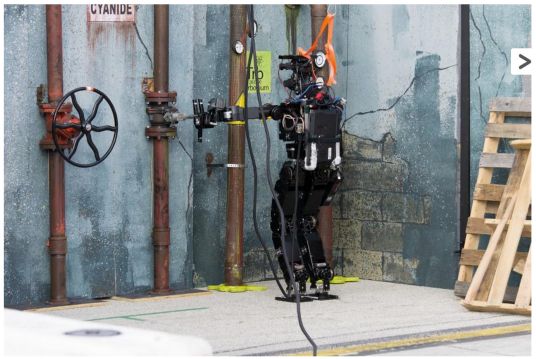
How: Reinforcement Learning

Reinforcement learning *is learning what to do – how to map situations to actions – so as to* ***maximize*** *a numerical* ***reward*** *signal in an* ***unknown uncertain*** *environment. The learner is not told which actions to take, as in most forms of machine learning, but she 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****).*

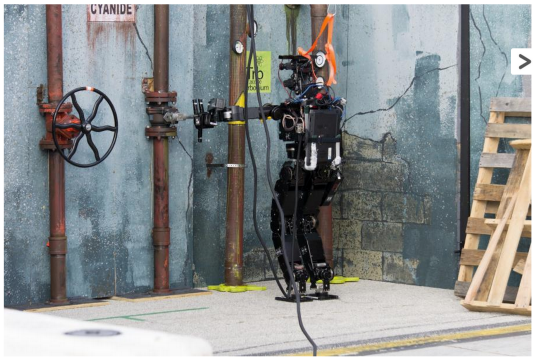
"An introduction to reinforcement learning",
Sutton and Barto (1998).

How: the Course

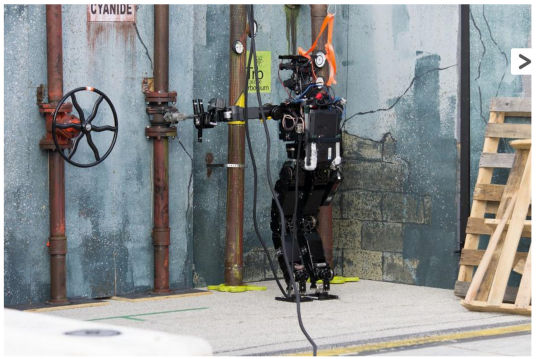
How: the Course



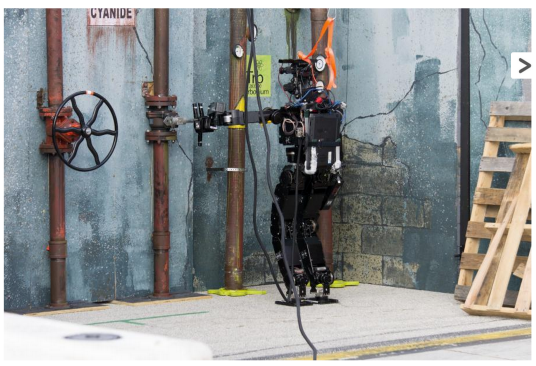
How: the Course

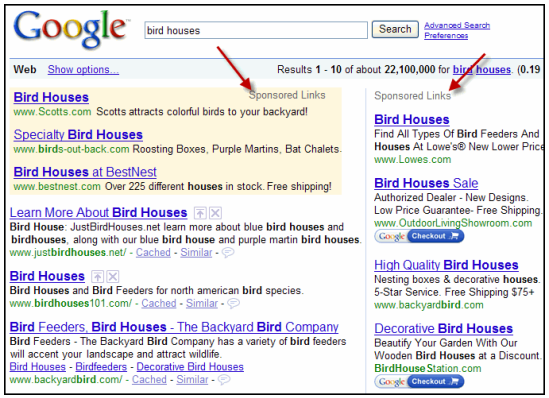


How: the Course

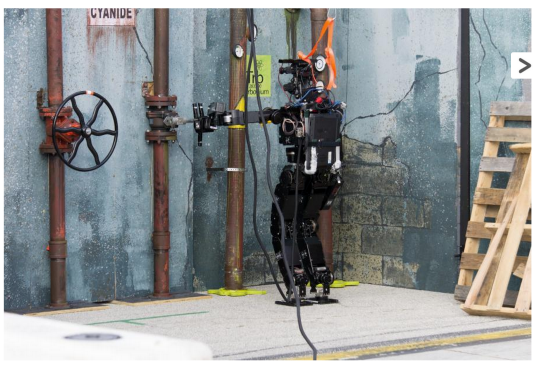


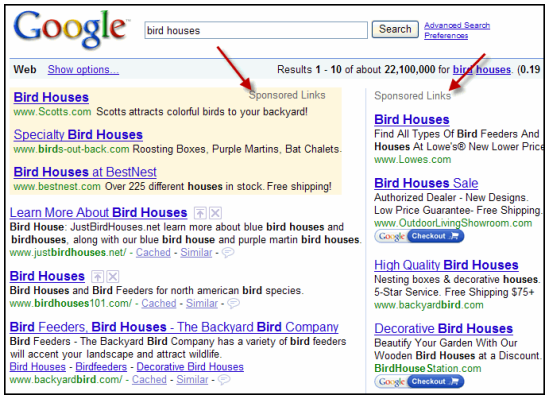
How: the Course



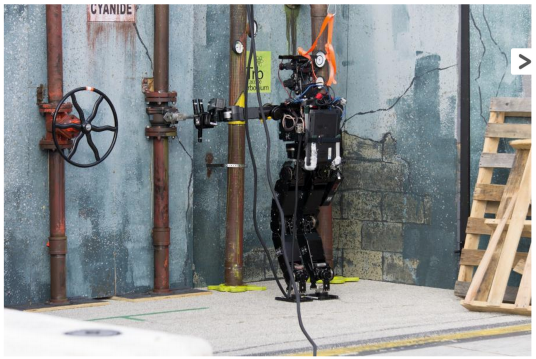


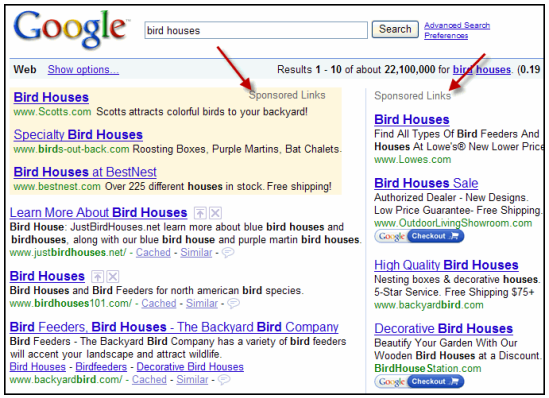
How: the Course

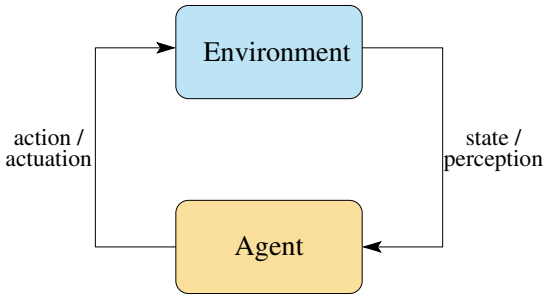




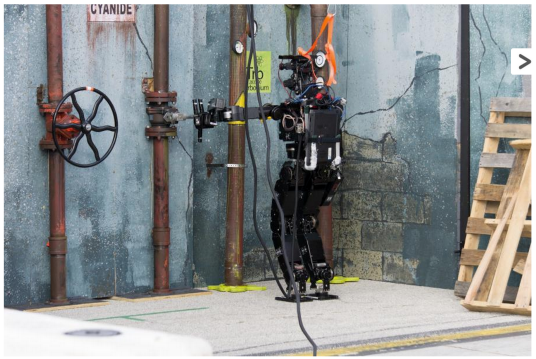
How: the Course

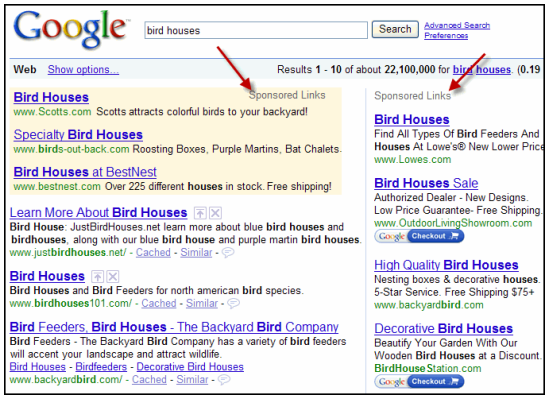


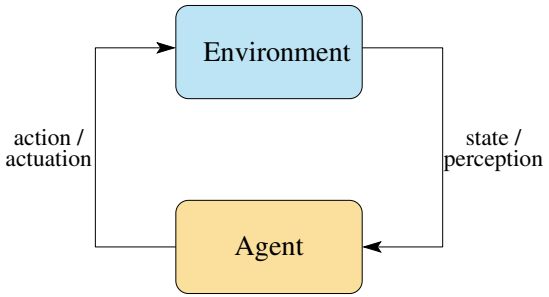




How: the Course







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

What: the Highlights of the Course

How 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 *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 *efficiently* explore in an RL problem

How to solve *approximately* an RL problem

- *What:* approximate dynamic programming
- *Tools:* statistical learning theory

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 *efficiently* explore in an RL problem

How to solve *approximately* an RL problem

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

When/What/Where

- 7 lectures
- 4 practical sessions (and homework) *[1 point each]*
- 1 final project (report and oral presentation) *[16 points]*

Opportunities for *spring internship* and *Ph.D.* positions.

When/What/Where

researchers.lille.inria.fr/~lazaric/Webpage/Teaching.html

Date	*Topic*	*Classroom*
29/09	Intro/MDP	Conference
06/10	Dynamic Programming	Condorcet
13/10	RL Algorithms	Condorcet
20/10	*TP* on DP and RL	Condorcet
27/10	Multi-arm Bandit (1)	Condorcet
03/11	*TP* on Bandit	Amphi Curie
10/11	Multi-arm Bandit (2) **[projects]**	Amphi Curie
17/11	*TP* on Bandit	Condorcet
24/11	Approximate DP	Condorcet
01/12	*TP* on ADP	Condorcet
08/12	Sample Complexity of ADP	Condorcet
15/12	Guest lecture	(TBD)
mid-Jan	Evaluation	(TBD)

Lectures are from 11am to 1pm, TP from 11am to 1:15pm.

Reinforcement Learning

Alessandro Lazaric

alessandro.lazaric@inria.fr

sequel.lille.inria.fr