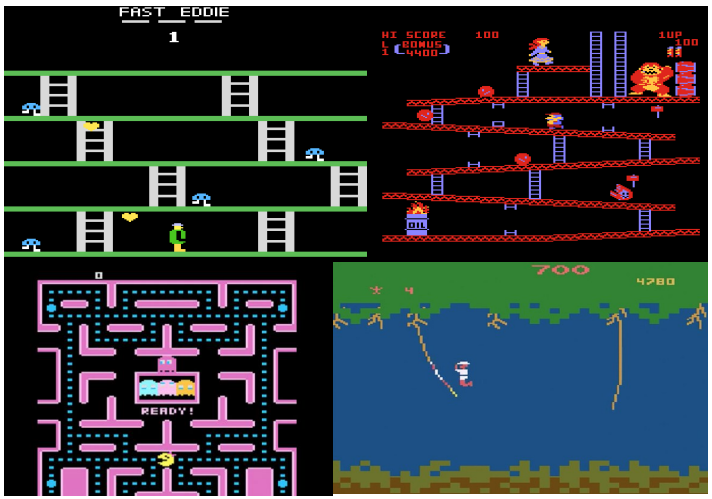
Reinforcement Learning

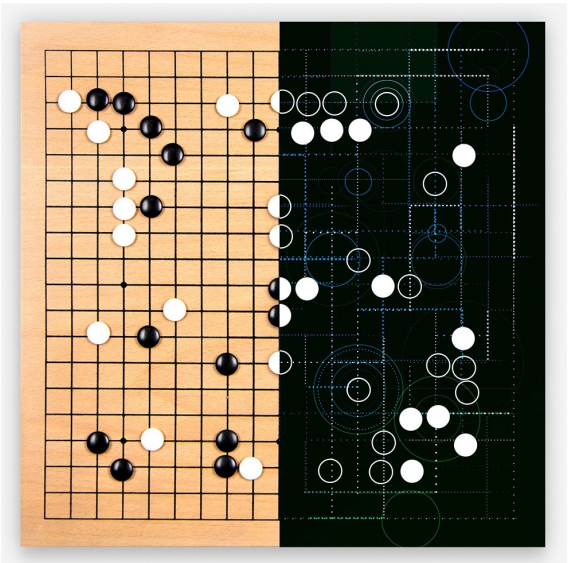
1. Introduction

Olivier Sigaud

Sorbonne Université
http://people.isir.upmc.fr/sigaud

Why this class (1)?



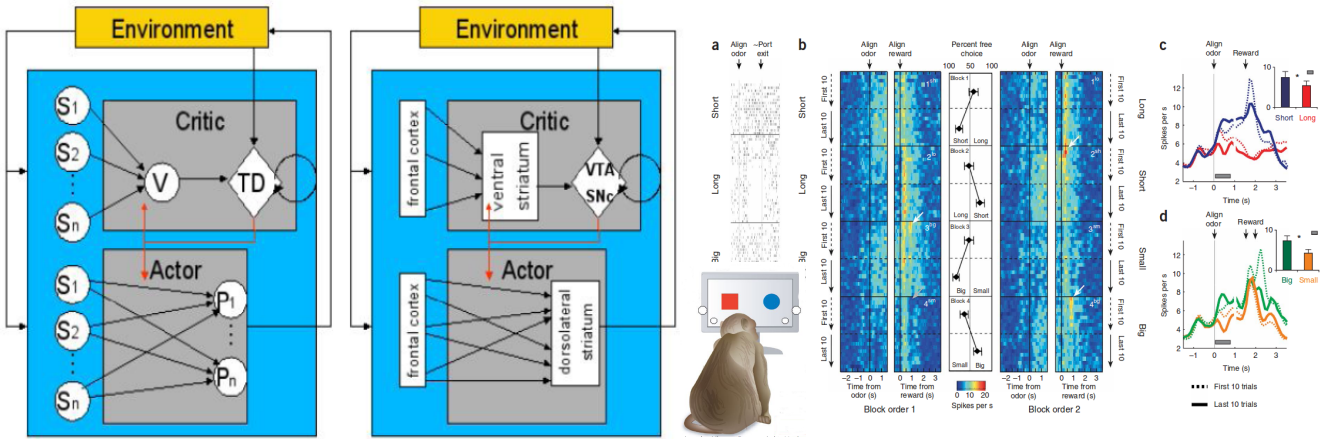


- A lot of buzz about deep reinforcement learning as an engineering tool

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015) Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. (2017) Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359

Why this class (2)?

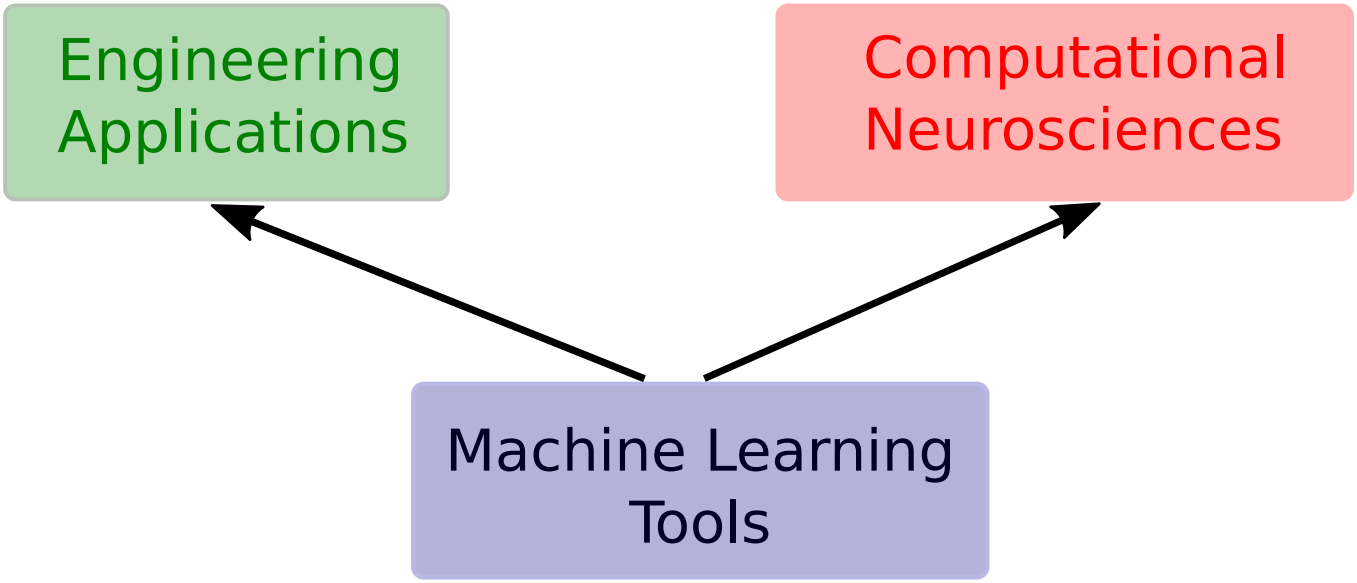


- The reinforcement learning framework is relevant in computational neuroscience
- This aspect will be left out

Takahashi, Y., Schoenbaum, G., & Niv, Y. (2008) Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. *Frontiers in neuroscience*, 2:14

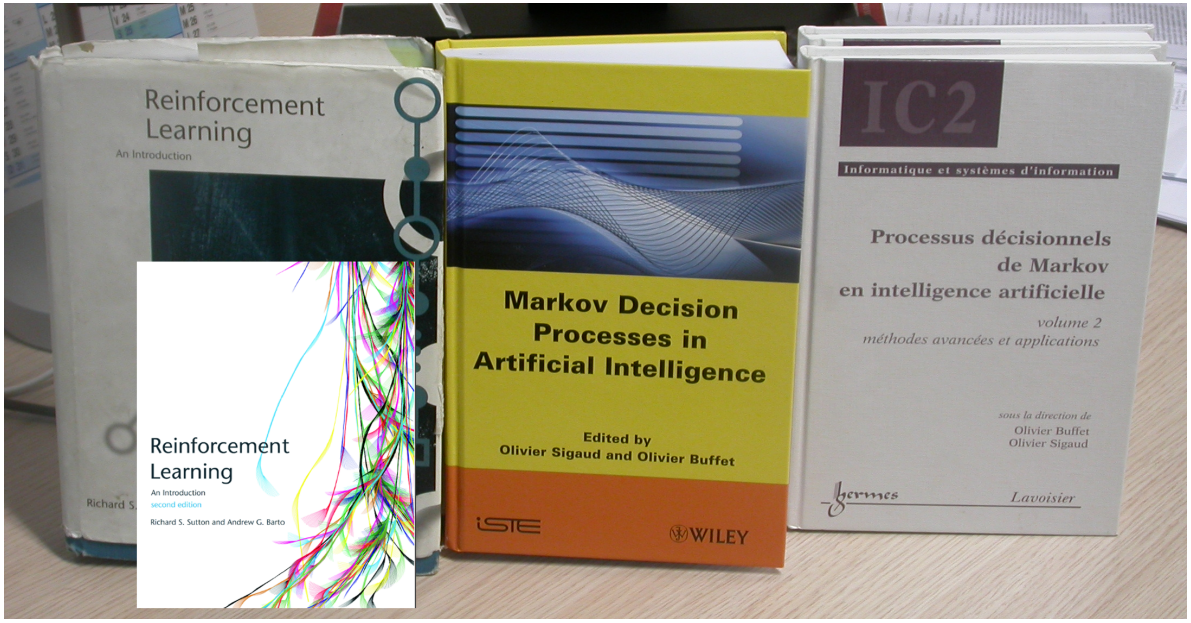
Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience*, 10(12):1615–1624

General goal of these videos



- Provide the machine learning background
- Use more videos than math
- Explain basic concepts from the discrete case
- Provide an introduction to deep RL algorithms
- And present the current big picture
- Related labs: `https://github.com/osigaud/rl_labs_notebooks.git`

Introductory books

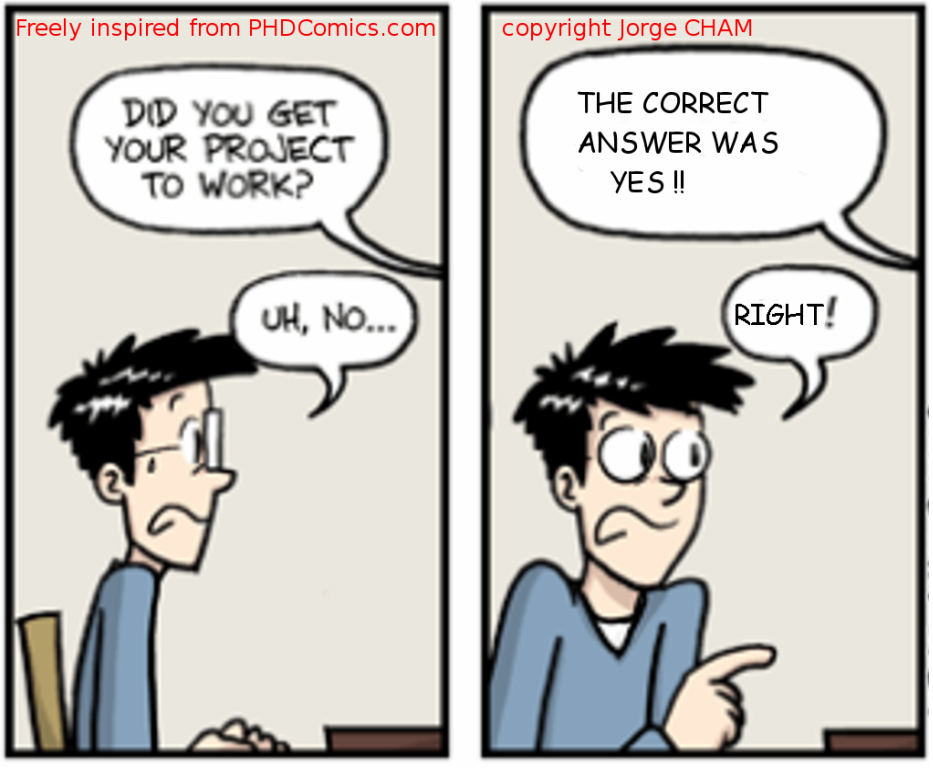


1. [Sutton & Barto, 1998]: the ultimate introduction to the field, in the discrete case
2. New edition available:

   `https://drive.google.com/file/d/1xeUDVGWGUUv1-ccUMAZHJLej2C7aAFWY/view`
3. [Buffet & Sigaud, 2008]: in french
4. [Sigaud & Buffet, 2010]: (improved) translation of 3

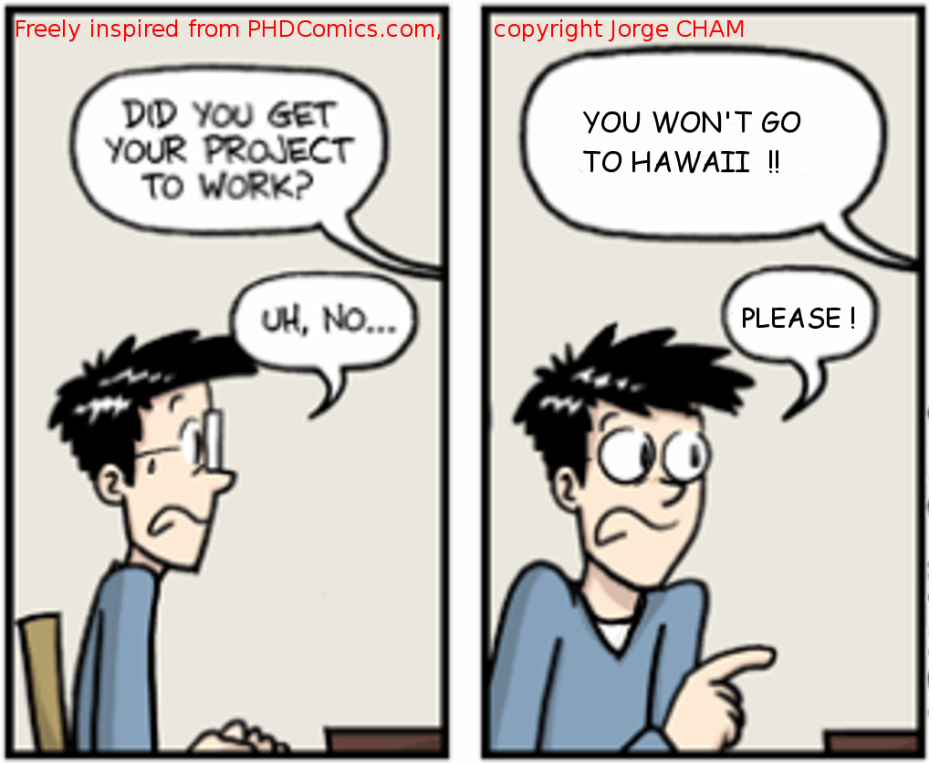
Sutton, R. S. & Barto, A. G. (1998) *Reinforcement Learning: An Introduction*. MIT Press.

Supervised learning



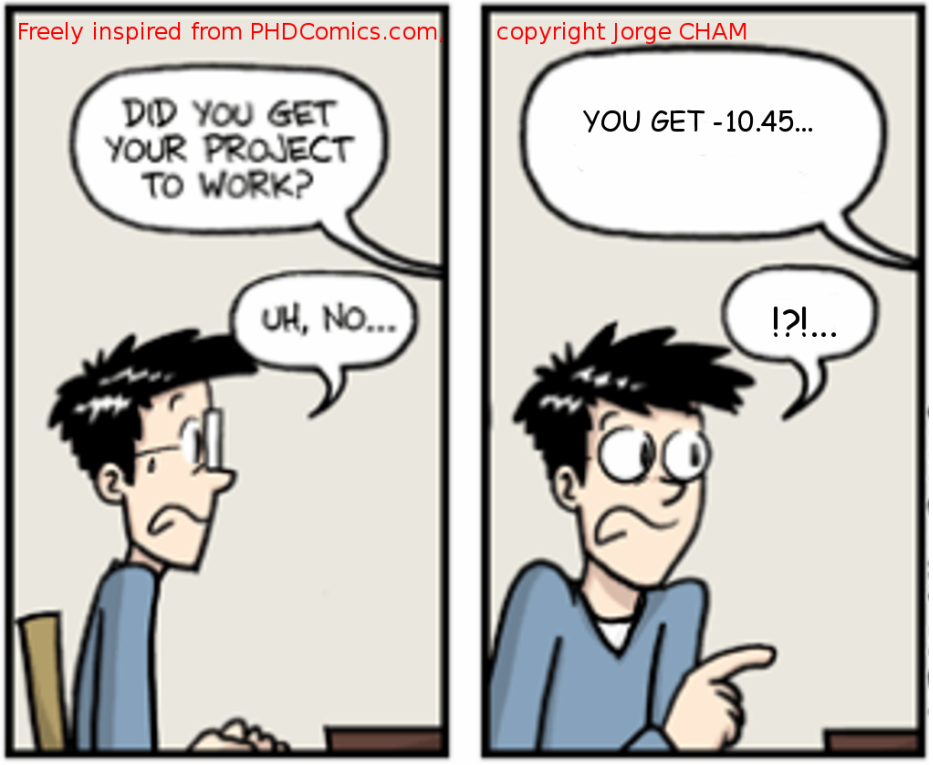
- The supervisor indicates to the agent the expected answer
- The agent corrects a model based on the answer
- Typical mechanism: gradient backpropagation, RLS
- Applications: classification, regression, function approximation...

Cost-Sensitive Learning



- The environment provides the value of action (reward, penalty)
- Application: behaviour optimization

Reinforcement learning



- In RL, the value signal is given as a scalar
- How good is -10.45?
- Necessity of exploration

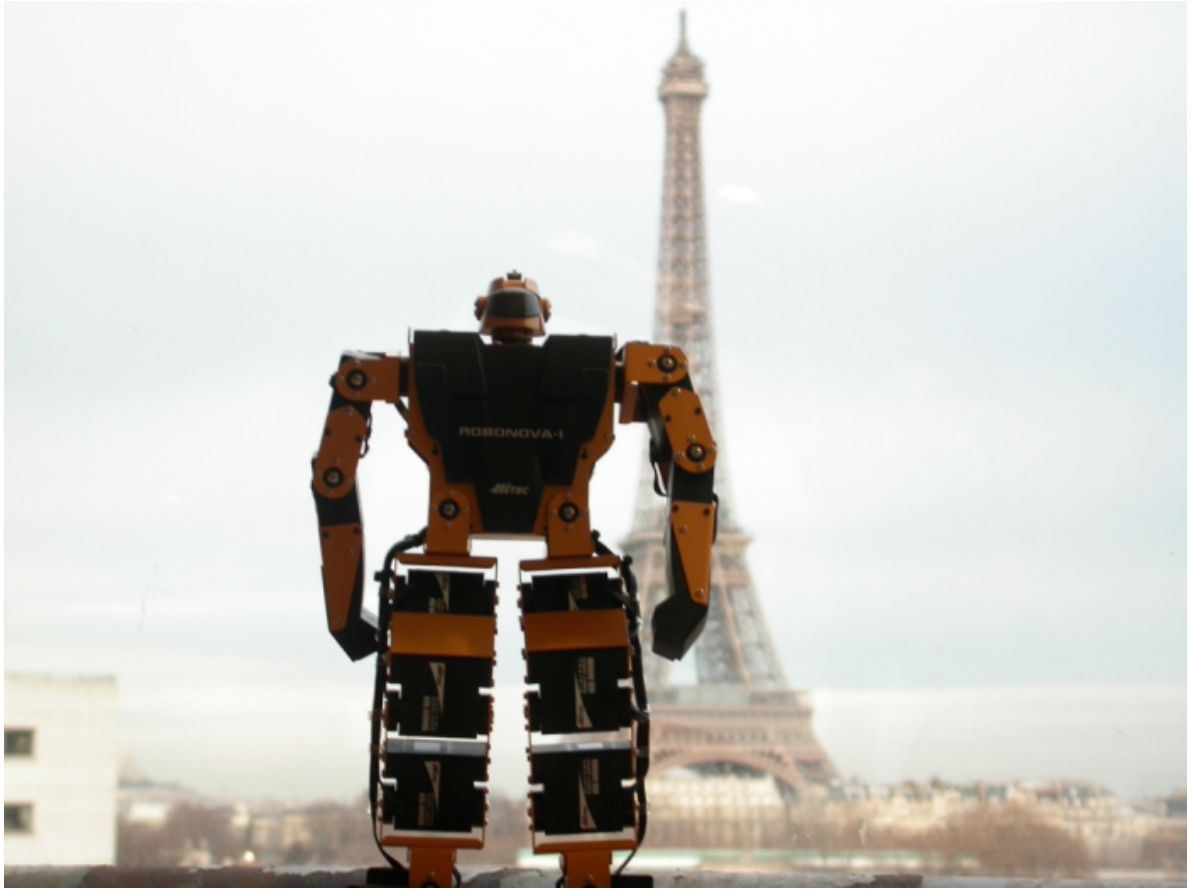
The exploration/exploitation trade-off

- Exploring can be (very) harmful
- Shall I exploit what I know or look for a better policy?
- Am I optimal? Shall I keep exploring or stop?
- Decrease the rate of exploration along time
- ϵ-*greedy*: take the best action most of the time, and a random action from time to time

Outline of next videos

1. Dynamic programming
2. Model-free Reinforcement Learning
3. Advanced discrete Reinforcement Learning
4. DQN
5. DDPG
6. Other deep RL research

Any question?



Send mail to: Olivier.Sigaud@upmc.fr

ISIR
INSTITUT
DES SYSTÈMES
INTELLIGENTS
ET DE ROBOTIQUE

Buffet, O. & Sigaud, O. (2008).

Processus décisionnels de Markov en intelligence artificielle.
Lavoisier.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015).
Human-level control through deep reinforcement learning.
Nature, 518(7540):529–533.

Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007).

Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards.
Nature Neuroscience, 10(12):1615–1624.

Sigaud, O. & Buffet, O. (2010).

Markov Decision Processes in Artificial Intelligence.
iSTE - Wiley.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. (2017).
Mastering the game of go without human knowledge.
Nature, 550(7676):354–359.

Sutton, R. S. & Barto, A. G. (1998).

Reinforcement Learning: An Introduction.
MIT Press.

Takahashi, Y., Schoenbaum, G., & Niv, Y. (2008).

Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model.
Frontiers in neuroscience, 2:14.