Course summary and final exam

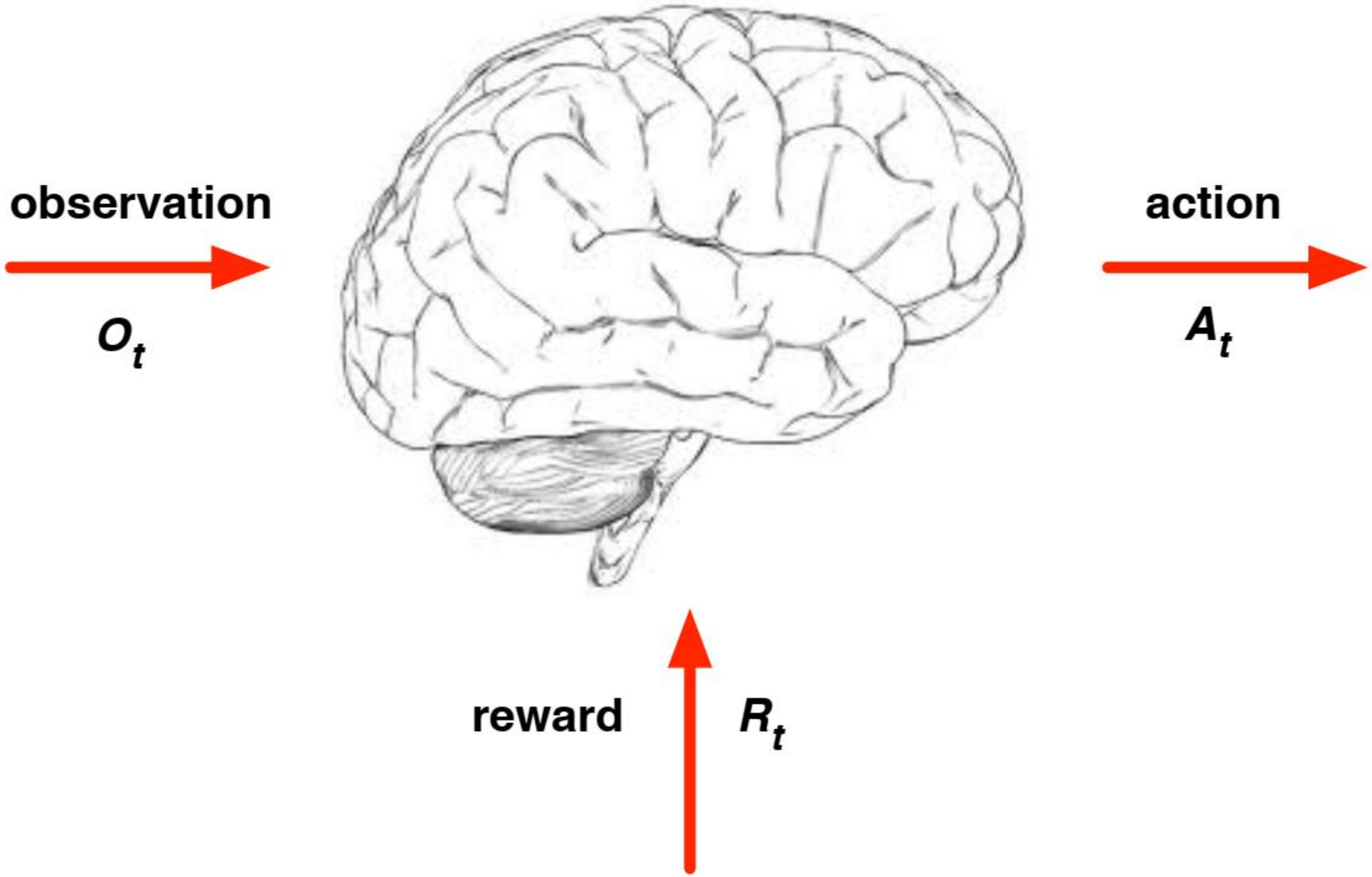
DAVIDE BACCIU – DAVIDE.BACCIU@DI.UNIPI.IT

UNIVERSITÀ DI PISA

Outline

- ✓ Course wrap-up
- ✓ An illustrated ex-post taxonomy of RL
- ✓ A sneak peak on research works
- ✓ Final projects and seminars

Course Wrap-Up



It had all started with this picture

Markov Decision Process

Definition (Markov Decision Process)

A Markov Reward Process is a tuple $\langle \mathcal{S}, \mathcal{A}, \boldsymbol{P}, \mathcal{R}, \gamma \rangle$

- $\mathcal{S}$ is a finite set of states
- $\mathcal{A}$ is a finite set of actions a
- $\boldsymbol{P}$ is a state transition matrix, s.t. $P_{ss'}^{a} = P(S_{t+1} = s' | S_t = s, A_t = a)$
- $\mathcal{R}$ is a reward function, s.t. $\mathcal{R}_s^a = \mathbb{E}[R_{t+1} | S_t = s, A_t = a]$
- γ is a discount factor, $\gamma \in [0,1]$

Key Components of an RL Agent…

An RL agent may be directly or indirectly trying to learn a

- ✓ Policy - Agent’s behaviour function
- ✓ Value function - How good is each state and/or action
- ✓ Model - Agent’s representation of the environment

$$S_1, A_1, R_2, \ldots, R_n, S_n$$

…and its *raison d'être*

Choose actions such as to maximize the discounted future reward (the return)

$$G_t = R_{t+1} + \gamma R_{t+2} + \cdots + \gamma^{T-1} R_T$$

3 Types of Reinforcement Learning (plus one)

Value-based

- ✓Learn the state or state-action value
- ✓Act by choosing best action in state
- ✓Exploration is a necessary add-on

Policy-based

- ✓Learn the stochastic policy function that maps state to action
- ✓Act by sampling policy
- ✓Exploration is baked in (trial-and-error)

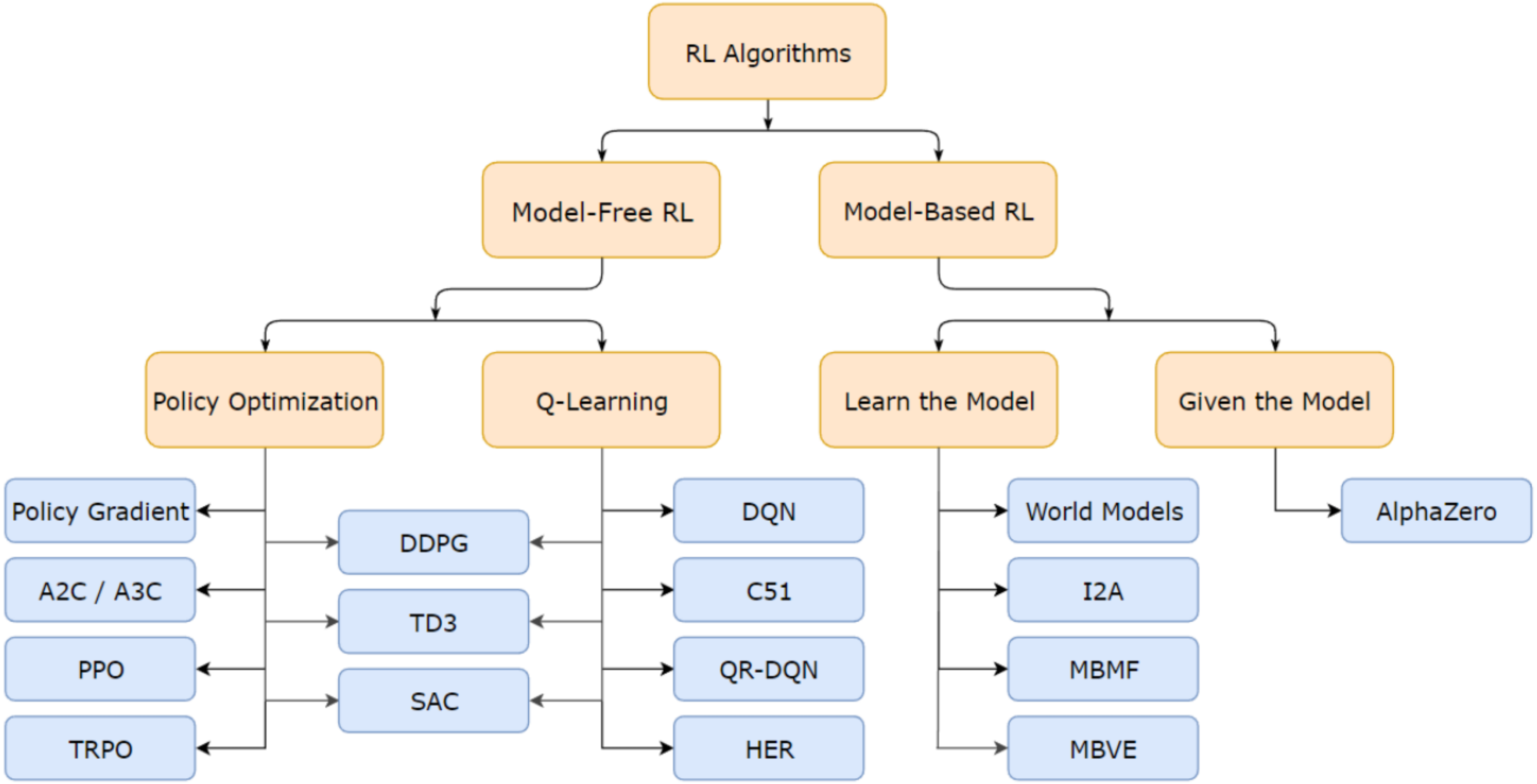
Model-based

- ✓Learn the model of the world, then plan using the model
- ✓Update model often
- ✓Re-plan often

Inverse Reinforcement Learning

- ✓Use demonstrations to learn a reward function and train a policy out of it
- ✓Does not need full demonstrations
- ✓Good generalization

A Taxonomy for RL



Exploration Vs Exploitation

- ✓Optimism in the face of uncertainty

$$a_t = \arg\max_{a \in \mathcal{A}} Q(a) + \sqrt{\frac{2 \log t}{N_t(a)}}$$

- ✓Despite what you could have expected, a hideout for heavily theoretical research

RL Research Sneak Peek

Other Topics We Did Barely Mention

- ✓ Inverse Reinforcement Learning
- ✓ Partial Observability
- ✓ Hierarchical reinforcement learning
- ✓ Curriculum learning
- ✓ Transfer Learning
- ✓ Meta-Learning
- ✓ Continual Learning
- ✓ Multi Agent Reinforcement Learning
- ✓ Scalable reinforcement learning

Some can be good candidates for final projects and seminars

Some Example Paper on the Research Topics

Wulfmeier et al, Maximum Entropy Deep Inverse Reinforcement Learning, 2015

Finn et al, Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization, ICML 2016

Igl et al, Deep Variational Reinforcement Learning for POMDPs, ICML 2018

Choi et al, Inverse Reinforcement Learning in Partially Observable Environments. JMLR 2011

Vezhnevets et al, FeUdal networks for hierarchical reinforcement learning, ICML 2017

Nachum et al, Data-Efficient Hierarchical Reinforcement Learning, NeurIPS 2018

Sainbayar Sukhbaatar, et al. “Intrinsic Motivation and Automatic Curricula via Asymmetric Self-Play.” ICLR 2018.

Carlos Florensa, et al. “Automatic Goal Generation for Reinforcement Learning Agents” ICML 2019.

Sebastien Racaniere & Andrew K. Lampinen, et al. “Automated Curriculum through Setter-Solver Interactions” ICLR 2020.

Allan Jabri, et al. “Unsupervised Curricula for Visual Meta-Reinforcement Learning” NeuriPS 2019.

Karol Hausman, et al. “Learning an Embedding Space for Transferable Robot Skills “ ICLR 2018.

Karl Cobbe, et al. “Quantifying Generalization in Reinforcement Learning” ICML 2019

Andrei A. Rusu et al. “Sim-to-Real Robot Learning from Pixels with Progressive Nets.” CoRL 2017.

Wojciech Marian Czarnecki, et al. “Mix & Match – Agent Curricula for Reinforcement Learning.” ICML 2018.

Seminars and Exams

M.Sc. Student Seminars

- ✓Student seminar lectures (n.2) as a conclusion of this course
  - ✓1 in July and 1 in September
  - ✓15 minutes presentation
  - ✓5 minutes Q&A on the content of the presentation
- ✓ Seminar content
  - ✓Read 3 relevant papers on a topic of interest for the course; summarize their content; confront the approaches providing your own considerations
  - ✓Implement 1 RL method from literature and attempt a validation on a simple application; describe the model, its implementation and the validation results
- ✓Seminar ideas: additional lectures on Moodle & research topics in this lectures
- ✓Everybody welcome to attend!

Ph.D. Students

- ✓ Read 3 (or more) relevant papers on a topic of interest for the course and summarize their content in a report (6-10 pages single column, NeurIPS format)
- ✓ Sketch/propose a novel RL method/application: report your idea in sufficient detail in a short paper (6-10 pages single column, NeurIPS format)
- ✓ Implement a RL-based application and validate it: prepare a short presentation to report the results (10-15 slides describing the model, the implementation and the results)
- ✓ Contact me and agree on alternative ways (e.g. using RL in your Ph.D. project)
- ✓ No presentation required, only delivery of material
- ✓ Please complete within the year 2021