Reinforcement Learning

Prof. Volkan Cevher
volkan.cevher@epfl.ch

Lecture 1: Introduction to Reinforcement Learning

Laboratory for Information and Inference Systems (LIONS)
École Polytechnique Fédérale de Lausanne (EPFL)

EE-568 (Spring 2024)

lions@epfl

License Information for Reinforcement Learning (EE-568)

Acknowledgements

◦ Joint effort between ODI@ETHZ (Prof. Niao He) and lions@epfl (myself)

- Anas Barakat, Batu Yardim, Jiawei Huang, Liang Zhang, Zebang Shen, Junchi Yang, Siqi Zhang, and Yifan Hu
- Luca Viano, Igor Krawczuk, Ali Kavis, Ahmet Alacaoglu, Grigorios Chrysos, Pedro Abranches, Leello Dadi, Ali Ramezani, Stratis Skoulakis, Kimon Antonakopoulos, Fanghui Liu, Fabian Latorre, Thomas Pethick, Angeliki Kamoutsi, Adrian Müller, Yongtao Wu, Wanyun Xie, Elias Abad Rocamora, Zhenyu Zhu, and Andrej Janchevski

Logistics

- **Credits:** 6
- **Lectures:** Thursdays 13:15-17:00 (GCC330, tentative)
- **Exercise hours:** Thursdays 13:15-17:00 (GCC330, tentative)
- **Prerequisites:** Previous coursework in optimization, probability theory, and linear algebra is required (i.e., EE-556 Math of Data). Familiarity with deep learning and programming in python is useful.
- **Grading:** Attendance (1pt), 3 Jupyter Notebooks (1pt each), Project (2pts) or Scribe (2pts, only PhD students may replace the project by writing lecture notes)
- **Moodle:** My courses > Genie electrique et electronique (EL) > Master
  > EE-568 Logistics & Course schedule & Learning materials
- **Details:** All details are explained in `https://go.epfl.ch/rl-moodle` (see Logistics & Course schedule)
- **TAs**: Luca Viano (Head TA), Adrian Müller (Co-Head TA), Leello Dadi, Yongtao Wu, Wanyun Xie, Elias Abad Rocamora, Zhenyu Zhu, Andrej Janchevski

Logistics for online teaching

- **Moodle:** `https://go.epfl.ch/rl-moodle`
- **Zoom:** `https://go.epfl.ch/rl-zoom`
- **Switchtube:** `https://go.epfl.ch/rl-lectures`

Project guidelines

- You can choose *one* of the following options:
  1. **Theory project:**
     - Read 3 theory papers in an active RL research area (we will provide pointers). Summarize them, understand which problems are still open (maybe solve them).
  2. **Practical project:**
     - Either implementing existing algorithms in new environments or try to improve existing algorithms on common environments.
     - Practical projects are in cooperation with EPFL labs. We will provide a list of labs you are allowed to reach out to.
  - In May, there will be a final report and a poster presentation of your project (1pt each).
  - You will work in groups of three people (both theory & practice). Registration opens on moodle soon.
- If and only if you are a PhD student, you may choose the following option instead:
  3. **Scribe:** Write lecture notes for a lecture assigned to you with a template we provide (2pts, no group work).
  - For each option 1–3, please check moodle for the detailed and binding guidelines.

A paradigm shift in machine learning (ML) applications

○ Self driving, industry automation, robotic manipulation, trading and finance,...

https://neptune.ai/blog/reinforcement-learning-applications

https://www.forbes.com/sites/bernardmarr/2022/12/28/what-does-chatgpt-really-mean-for-businesses/?sh=27bc344f7d1e

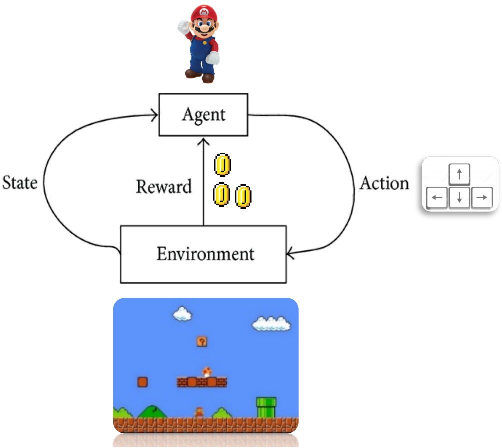
What is reinforcement learning (RL)?

◦ Classical definitions:

- ▷ Sutton and Barto, 1998: Reinforcement learning is learning what to do – how to map situations to actions – so as to maximize a numerical reward signal.
- ▷ Wikipedia, 2023: Reinforcement learning is an area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward.

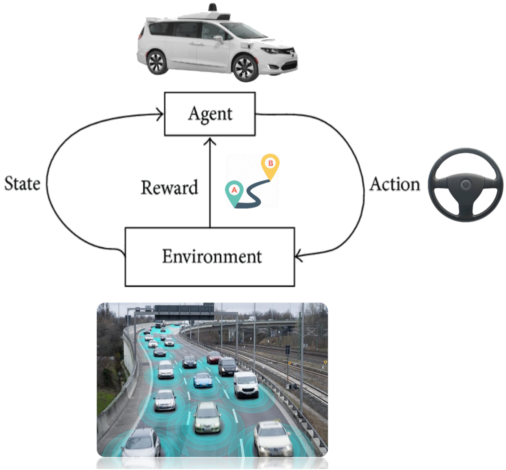
A common theme in RL

◦ An agent learns to act by interacting with an uncertain environment



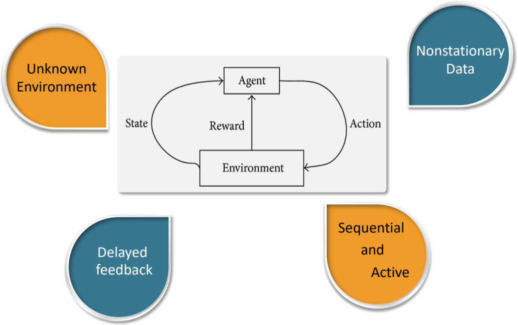
A common theme in RL

- An agent learns to act by interacting with an uncertain environment



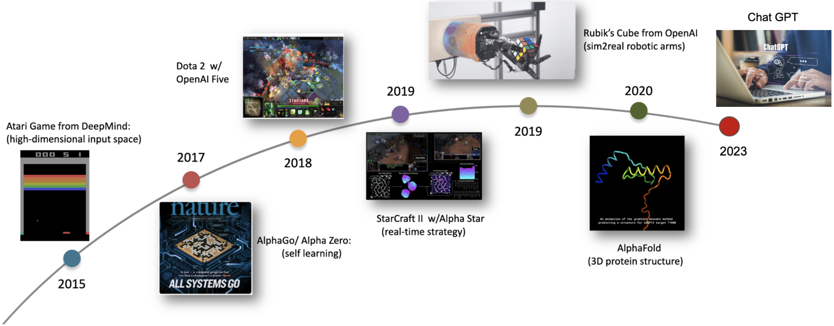
A common theme in RL

- An agent learns to act by interacting with an uncertain environment



Remarkable progress on ML applications

○ Which one is not due to RL?



Perceptions of RL

◦ Roughly speaking...

- For EE, it is control theory *mutatis mutandis*:

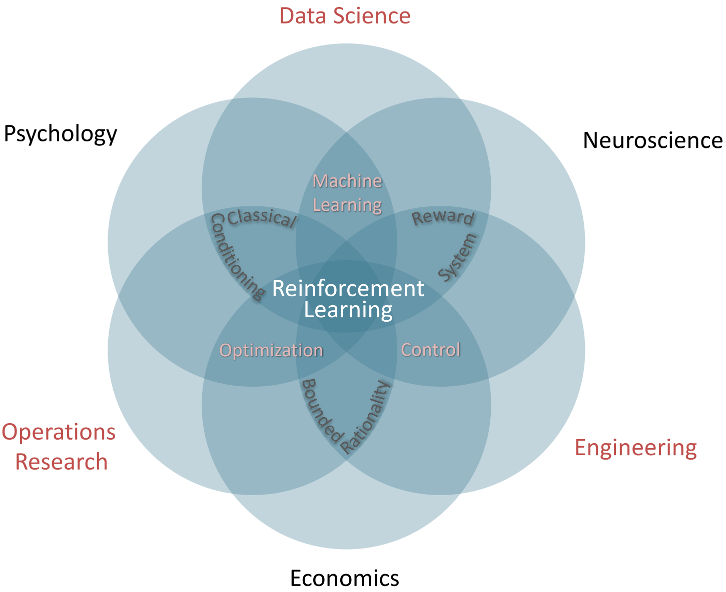
  control $\rightarrow$ action

  controller $\rightarrow$ agent or policy

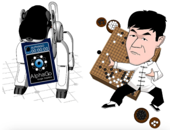
  system or plant $\rightarrow$ environment

- For CS, it is an ML paradigm along with supervised and unsupervised learning.

- For others?


Data Science
Psychology
Neuroscience
Machine Learning
Classical Conditioning
Reward System
Reinforcement Learning
Optimization
Control
Bounded Rationality
Operations Research
Engineering
Economics

Challenges to RL



The New York Times

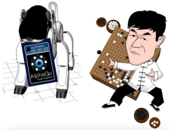
One Giant Step for a Chess-Playing Machine

The stunning success of AlphaZero, a deep-learning algorithm, heralds a new age of insight — one that, for humans, may not last long.

What is frustrating about machine learning, however, is that the algorithms can't articulate what they're thinking. We don't know why they work, so we don't know if they can be trusted. AlphaZero gives every appearance of having discovered some important principles about chess, but it can't share that understanding with us. Not yet, at least. As human beings, we want more than answers. We want insight. This is going to be a source of tension in our interactions with computers from now on.

- Theoretical foundations are more important than ever.

Challenges to RL



The New York Times

One Giant Step for a Chess-Playing Machine

The stunning success of AlphaZero, a deep-learning algorithm, heralds a new age of insight — one that, for humans, may not last long.

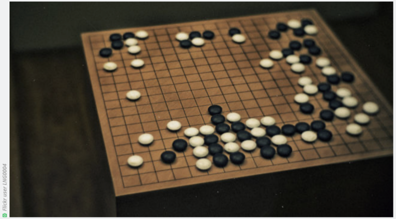
What is frustrating about machine learning, however, is that the algorithms can't articulate what they're thinking. We don't know why they work, so we don't know if they can be trusted. AlphaZero gives every appearance of having discovered some important principles about chess, but it can't share that understanding with us. Not yet, at least. As human beings, we want more than answers. We want insight. This is going to be a source of tension in our interactions with computers from now on.

DEEP BLUE WAS JUST THE START —

Man beats machine at Go in human victory over AI

Amateur exploited weakness in systems that have otherwise dominated grandmasters.

RICHARD WATERS, FINANCIAL TIMES - 2/19/2023, 1:51 PM



- Theoretical foundations are more important than ever.
- Common challenges with ML: Robustness, interpretability, scalability, reproducibility, etc.

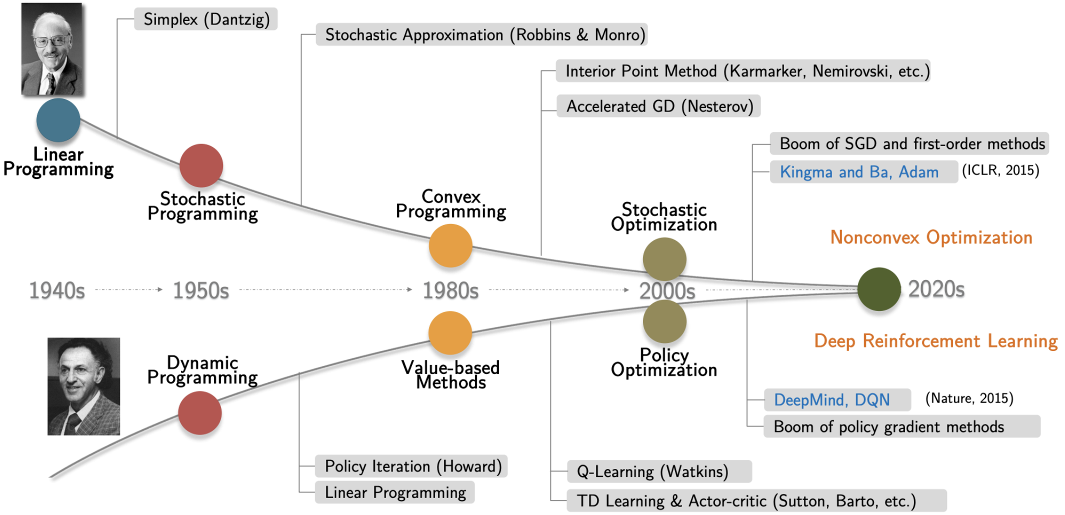
Perceptions of our RL course (EE-568)

- Why are you taking this course?
  - ▷ Learn the basics of RL
  - ▷ Gain hands-on experience with RL implementations
  - ▷ Apply RL to my research
  - ▷ Might be useful for my future job
  - ▷ Just need the credits
  - ▷ Other reasons?
  - ▷ Let us know `https://go.epfl.ch/rl_form2024`

What are our learning objectives?

- By the end of the course, participants will be able to
  - ▷ Define the key features of RL that distinguishes it from standard ML
  - ▷ Identify the strengths and limitations of various RL algorithms
  - ▷ Understand the theoretical properties of RL algorithms
  - ▷ Recognize the common, connecting boundary of optimization and RL
  - ▷ Formulate and solve sequential decision-making problems by applying relevant RL tools
  - ▷ Generalize or discover "new" applications, algorithms, or theories of RL towards conducting research

What EE-568 is really about: Theory and methods



What EE-568 is *not* really about: Product development

- The following important topics are beyond the scope of this course:
  - Coding tricks and super practical implementations of RL
  - Product development of RL in real-world
  - RL engineering
  - Physically building autonomous robots
  - Physically building autonomous driving systems
  - Building ChatGPT-like systems with RLHF

Should I take this course?

- This course is right for you, if you
  - want to understand the RL foundations
  - have interest in performing RL research
  - have strong math background
  - want to gain hands-on experience with RL

- This course may not be right for you, if you
  - **only** want to gain hands-on experience with RL
  - are not interested in formulating RL in applications
  - have only basic math background
  - want to develop deep learning expertise

What is left for the course?

- A preview of the course
  - Dynamic Programming
    - Value Iteration
    - Policy Iteration
    - Monte Carlo Methods
    - TD, SARSA, Q-learning
  - Linear Programming
    - Primal-Dual RL, REPS, Proximal Point
    - Applications to offline RL
  - Policy-based RL
    - Policy Gradient Method
    - Natural Policy Gradient Method
    - TRPO and PPO
  - Imitation Learning and Inverse RL
    - Behavior Cloning, GAIL
    - Interactive IL (DAgger, SMILe)
    - Max Margin and Max Entropy IRL
  - Markov Games
    - Fictitious Play
    - Policy Gradient
    - Nash Q-learning
  - Robust and Deep RL
    - Deep Q Network and Extensions
    - Deep Actor-Critic (A3C, DDPG, TD3)
    - Robust DDPG/TD3

Theory	Bellman Equations Policy Gradient Theorems Performance Difference Lemma	Stochastic Approximation Optimization and Game Theory Convergence Analysis

What's beyond?

- Episodic RL
- Strategic Exploration in RL
- Batch and Offline RL
- Safety in RL
- Multi-task RL
- Preference-based RL
- Causal RL
- Partially Observable Markov Decision Process (POMDP)
- ....

Where to go from here?

- Upcoming events:
  - **Multi-Agent RL Summer School @ EPFL, organized by our lab**
    https://edu.epfl.ch/coursebook/en/multi-agent-reinforcement-learning-EE-806
  - European Workshop on Reinforcement Learning, 2024
    https://ewrl.wordpress.com/ewrl17-2024/
  - RL Conference, 2024
    https://rl-conference.cc
  - Workshops at NeurIPS, ICML, ICLR, AAMAS, etc. (TBA)

- Recent workshops:
  - Goal-Conditioned Reinforcement Learning Workshop, NeurIPS 2023
    https://neurips.cc/virtual/2023/workshop/66519
  - Agent Learning in Open-Endedness Workshop
    https://neurips.cc/virtual/2023/workshop/66527
  - New Frontiers in Learning, Control, and Dynamical Systems Workshop, ICML 2023
    https://icml.cc/virtual/2023/workshop/21470
  - Reincarnating Reinforcement Learning Workshop, ICLR 2023
    https://iclr.cc/virtual/2023/workshop/12833
  - AI for Agent-Based Modelling (AI4ABM) Workshop, ICLR 2023
    https://iclr.cc/virtual/2023/workshop/12840
  - European Workshop on Reinforcement Learning, 2023
    https://ewrl.wordpress.com/past-ewrl/ewrl16-2023/

Where to go from here? (cont'd)

- Seminars:
  - RL Theory Virtual Seminar:
    https://sites.google.com/view/rltheoryseminars/
  - Simons Institute Theory of RL:
    https://simons.berkeley.edu/programs/theory-reinforcement-learning
  - Simons Institute Learning and Games:
    https://simons.berkeley.edu/programs/games2022/workshops#simons-tabs

Questions

That's it! Any questions?