Foundations of Reinforcement Learning

From Markov Decision Processes to Optimal Value Functions

Matthia Sabatelli

October 13, 2021

Today's Agenda

Course Information

- Coordinator: Matthia Sabatelli
- Lecturers: Matthia Sabatelli (m.sabatelli@rug.nl) and Nicole Orzan (n.orzan@rug.nl)
- Classroom: TBD
- Theoretical Lectures: Monday morning from 9:00-11:00
- Computer Labs: Monday afternoon from 15:00-17:00

Course Information

- Lecture 1: Foundations of Reinforcement Learning (Matthia)
- Lecture 2: Exploration and Bandit Problems (Nicole)
- Lecture 3: Dynamic Programming (Nicole)
- Lecture 4: Model-Free Reinforcement Learning (Matthia)
- Lecture 5: Function Approximators (Matthia)
- Lecture 6: Beyond Model-Free Reinforcement Learning (Matthia)
- Lecture 7: *What does it mean to do research in RL?* (Matthia & Nicole)

Course Information

All course material will be made available

- Nestor
- Github: `https://github.com/paintception/reinforcement-learning-practical`

Course Information

Textbook: Reinforcement Learning: An Introduction by Sutton & Barto

Course Information

Final course assessement:

- There is **no exam**
- Students should handle in three deliverables:
  1. Assignment 1: 25% of the grade (coding)
  2. Assignment 2: 25% of the grade (mathematics)
  3. Report: 50% of the grade (final project)
- Students can work alone or in groups of a maximum of **2** people

Reinforcement Learning

Machine Learning is typically divided into three branches:

1. Supervised Learning: learning from labeled data
2. Unsupervised Learning: learning from unlabeled data
3. Reinforcement Learning: learning from experience

Reinforcement Learning

A reminder of supervised learning:

- input space: $\mathcal{X}$
- output space: $\mathcal{Y}$
- probability distribution $p(x, y)$

The goal

We want to build a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ that minimizes the expectation of a given loss ℓ

$$\mathbb{E}_{(x,y)\sim p(x,y)}\{\ell(y, f(x))\}$$

through leaning samples $LS = \{(x_i, y_i)|i = 1, ..., N\}$ of input-ouput pairs drawn from $p(x, y)$.

Reinforcement Learning

Supervised learning is therefore:

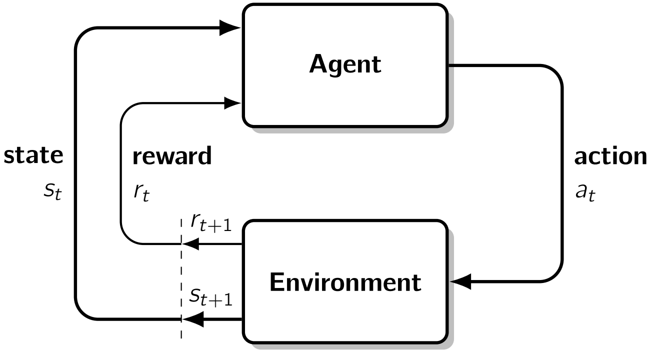
- a learning paradigm which is static
- no interaction happens between the learner and $p(x, y)$
- Assumes we have access to a knowledgable supervisor

In Reinforcement Learning however ...

Reinforcement Learning

- We would like to learn how to interact with an environment
- We do not assume any sort of supervision but only a reward signal
- The component of time plays a crucial role
- The learning process is therefore dynamic and uncertain

Reinforcement Learning



Markov Decision Processes

Value Function