

Reinforcement Learning for Markov Decision Processes

Introduction

Idris Kharroubi

LPSM
Sorbonne Université

DUFE

Outline

- 1 Goal
- 2 Why solving such a problem?
- 3 How to proceed
- 4 Table of contents

Applying the theory of optimal control

- Main goal : find a way to apply the theory of optimal control for daily problems
- First simplification: consider a discrete time framework.
- Not enough: In real life, we do not know the model describing the controlled dynamics.
- Need to set numerical methods that allow to control and learn the model at the same time.

Outline

- 1 Goal
- 2 Why solving such a problem?
- 3 How to proceed
- 4 Table of contents

Applications in finance

- Portfolio management
- Market making
- LOB

Outline

- 1 Goal
- 2 Why solving such a problem?
- 3 How to proceed
- 4 Table of contents

- First present the control problem.

Rich literature beginning from the late 50' with two main approaches :

Dynamic programming of Bellman	Vs	Maximum principle of Pontryaguine
-----------------------------------	----	--------------------------------------

RL adopt the Bellman DP $\Rightarrow$ present the theory for MDP

- Use the DP approach to derive numerical approximations
- Need also to consider **randomized strategies** to allow for exploration and learning.

Outline

- 1 Goal
- 2 Why solving such a problem?
- 3 How to proceed
- 4 Table of contents**

Ch. 1 Optimal control of MDP

Ch. 2 Learning MDP

Ch. 3 Value-based algorithms

Ch. 4 Policy-based algorithms