University of Mannheim

Bachelor Thesis

Markov-Decision Processes

by

Felix Benning

born on the 27.11.1996 in Nürtingen
matriculation number 1501817

in the

Fakulty for Mathematics in Business and Economics

Supervisor: Prof. Dr. Leif Döring

Due Date: ???

Declaration of Authorship

I hereby declare that the thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references.

This thesis was not previously presented to another examination board and has not been published.

City, Date   Signature

Preface

Contents

Introduction

Chapter 1

Markov Decision Processes

Definition 1.0.1. (Kernel) $(X, \mathcal{A}_Y), (X, \mathcal{A}_X)$ measure spaces
$\lambda\colon \mathcal{X} \times \mathcal{A}_Y \to \mathbb{R}$ is a *kernel* $:\iff$ $x \mapsto \lambda(x, A)$ measurable
$A \mapsto \lambda(x, A)$ a measure

Definition 1.0.2. (Markov Decision Process - MDP)
$\mathcal{M} = (\mathcal{X}, \mathcal{A}, \mathcal{P}_0)$, with:

$\mathcal{X}$ countable (finite) set of states
$\mathcal{A}$ countable (finite) set of actions

$$\mathcal{P}_0\colon \begin{cases} \mathcal{X} \times \mathcal{A} \to \not{P}(\mathcal{X} \times \mathbb{R}) \\ (x, a) \mapsto \mathbb{P}(\cdot \mid x, a) \end{cases}$$

$\not{P}(\mathcal{X} \times \mathbb{R})$ the set of probability measures on $\mathcal{X} \times \mathbb{R}$,
$\mathcal{X}$ represents the next states,
$\mathbb{R}$ the payoffs

is a *(finite) Markov Decision Process*
Together with a discount factor $\gamma \in (0, 1]$ it is a:
discounted reward MDP $\quad \gamma < 1$
undiscounted reward MDP $\quad \gamma = 1$
For $(Y_{(x,a)}, R_{(x,a)}) \sim \mathcal{P}_0(\cdot \mid x, a)$ a random variable, is

$$r(x, a) \coloneqq \mathbb{E}[R_{(x,a)}] \quad \text{the } \textit{immediate reward function}$$

An MDP is *evaluated* as follows:

Chapter 2

Title Chapter 2

Bibliography