Markov Decision Processes:

This version: (First official draft)

Course instructors
Name: G.P. Kiesmüller
Mail: office.cdt@mgt.tum.de

Location
Campus Heilbronn

Application and registration procedure

Prerequisites
Participants should have basic knowledge in probability theory and Markov Chains (Random variables, discrete and continuous distribution functions, conditional distributions, moments of random variables, discrete time Markov Chains, Continuous Time Markov Chains). Additionally, students should have basic programming experience.

Application/Registration process
There is a limit of 15 participants maximum for this course. Please use the official application procedure for the TUM MGT Doctoral Summer School.

The application starts on June 3, 2024, 12:00 noon. The deadline for application is June 30, 2024, end of day. Registration works on a first come, first served basis. Registration is only possible for doctoral candidates of TUM School of Management.

Please find the link for registration here: https://www.mgt.tum.de/faculty-research/doctoral-program/course-program

Course aims and objectives

Many real life systems are subject to uncertainty and should therefore be modelled with stochastic models. In this course, we focus on the theory and the application of Markov Decision Processes and Semi Markov Decision Processes. The students should gain knowledge about these models such that they are able to construct these models and apply them to solve real life problems. For illustration, we use among others, models of inventory systems, manufacturing systems and maintenance systems. We practice to derive the Bellmann equation for these systems and show how an optimal solution can be computed numerically. Besides the traditional solution approaches, we also discuss approaches based on reinforcement learning.

Content
We first define all elements of a discrete time Markov Decision Process and show how dynamic systems can be modelled with the Bellman equation. We discuss the existence of an
optimal stationary policy and show how it can be computed numerically. We distinguish between the expected total discounted reward criteria and long-run average costs. We discuss policy iteration, value iteration, linear programming, and reinforcement learning algorithms to derive optimal stationary policies.

**Preliminary schedule**
The course takes place in the second week of the TUM MGT Doctoral Summer School from September 16 to September 20, 2024. Please refer to the schedule for the Summer School for further details on the course schedule. The schedule for the Summer School can be found in the digital flyer on the Summer School: [https://www.mgt.tum.de/faculty-research/doctoral-program/course-program](https://www.mgt.tum.de/faculty-research/doctoral-program/course-program)

**Core readings**

**Assessment**
Students have to participate actively in the course and especially in the exercise sessions. They have to present the solutions of the exercises in the class.

Students have to bring their own laptop to the course and must be able to implement the studied algorithms (for example in R or in Python) to obtain solutions to the problems.

**Workload**
3 ECTS (21 hours lectures including exercise sessions)