Doctoral School in Finance and Economics

UNIVERSITÉ DU LUXEMBOURG

Course ID Decision Making Over Time Under Uncertainty: Markov Decision Processes

1. Course details

Semester:	2
Credit rating:	2
Pre-requisite(s):	Probability Theory (Poisson process, Markovian (reward) processes)
Lecturer(s):	Joachim Arts (joachim.arts@uni.lu)
Administrator:	Roswitha Glorieux
Tutor(s):	
Seminar times and rooms:	Block/Semester Course
Tutorial times and rooms:	tbd
Communications	**It is important that students should regularly read their University e-mails, as important information will normally be communicated this way.**
Mode of assessment:	Grading: homeworks/assignments 100% (3 homework assignments)
Examination Periods:	
Course WebPage:	Moodle.uni.lu

2. Aims and objectives

Aims
This course introduces students to the theory of Markov decision processes (MDPs). These processes are the archetypical model for making decisions over time under uncertainty. Decisions in supply chains and logistics must often be made under considerable uncertainty that will unfold over time. This course will equip students to model such decision problems as Markov decision processes and analyze them. In particular, they will study computational techniques to find numerical answers to decisions problems and analytical techniques to study structural properties of optimal decision rules. We will also give a brief outline of reinforcement learning.
Learning Objectives The student who followed this course - Can formulate a decision problem and correctly identify the state-space, action-space, transition laws and cost structure - Can use value iteration and linear programming to numerically solve MDPs on a finite and infinite time horizon. - Understands how reinforcement learning is an approximate numerical solution technique for MDPs - Can prove structural properties of optimal policies by induction proofs on the value function. In particular, they can use the event-based dynamic programming framework. - Can read literature in scientific journals that use MDPs

3. Plan of semester

tbd

LCL room B28

Session 1	14-Feb	13:00-16:00
Session 2	21-Feb	13:00-16:00
Session 3	28-Feb	13:00-16:00
Session 4	08-Mar	13:00-16:00
Session 5	14-Mar	13:00-16:00
Session 6	22-Mar	13:00-16:00
Session 7	29-Mar	13:00-16:00
Session 8	21-Apr	13:00-16:00
Session 9	26-Apr	13:00-16:00
Session 10	03-May	13:00-16:00

4. Course details (by topics)

- Finite state, finite actions space, discrete time finite horizon Markov Decision processes (Value iteration and formulation)
- Finite state, finite action space infinite horizon Markov Decision processes: Average and discounted reward criteria.

- Value iteration and Linear programming algorithms
- Continuous time MDPs and uniformization
- Value function properties: Monotonicity, Convexity, Sub-modularity, Super-Convexity, K-convexity.
- Operator based dynamic programming
- Induction proofs of structural properties of MDPs

5. Reference list/ Bibliography

The content of the course will be mostly based on:

- Markov Decision Processes Lecture Notes (by Lodewijk Kallenberg)
- Monotonicity in Markov reward and decision chains: Theory and applications. Foundations and Trends in Stochastic Systems 1:1-76, 2006.
- Stochastic Optimization Lecture note by Ger Koole and Sandjai Bhulai
- Tijms H. (2003) Stochastic Models, Wiley
-

Additional reading material:

- Porteus, E.L. (2002) Foundations of Stochastic Inventory Theory, Stanford Business Books
- Kulkarni, V.G. (1999) Modeling Analysis Design and Control of Stochastic Systems, Springer
- Puterman, M.L. (1994) Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley-Interscience.

6. Further information about assessment

Students will receive three homework assignments and the grade will be weighted average of these three assignments. In these assignments, students need to prove results for slight variations of models and they need to implement computational models.