

# MIE1615: MARKOV DECISION PROCESSES

---

DEPARTMENT OF MECHANICAL AND INDUSTRIAL ENGINEERING, UNIVERSITY OF TORONTO

- Reference:** “Markov Decision Processes - Discrete Stochastic Dynamic Programming”, Martin L. Puterman, Wiley, 1994.  
“Approximate Dynamic Programming”, Warren Powell, Wiley, 2007.  
“Neuro-Dynamic Programming”, Dimitri Bertsekas and John Tsitsiklis, Athena Scientific, 1996.
- Instructor:** Chi-Guhn Lee, MC322, 946-7867, cglee@mie.utoronto.ca.
- Office Hours:** By appointment.
- Course Description:** This is a course to introduce the students to theories and applications of Markov decision processes. Emphasis will be on the rigorous mathematical treatment of the theory of Markov decision processes. Topics will include MDP finite horizon, MDP with infinite horizon, and some of the recent development of solution method.
- Course Assessment:** **Assignments (20%):** There will be two assignments.  
**Exams (40%):** There will be one two-hour exam in class on Mar/11.  
**Project (40%):** Each group of at least **three** and up to **five** students will be asked to give three 25-minute lectures: one from each of Theory, RL and Application collections.
- Important Dates:** **Feb/11:** No lecture.  
**Mar/11:** Exam in class  
**Apr/8:** No lecture, student presentation only.
- Policies:** Late submission of assignments will NOT be accepted.  
Requests for regrading of assignments or exams will be considered only within one week from the time the graded work is returned in class.
- Project Presentation Schedule:** **Feb/25:** T1, T2, T3  
**Mar/4:** T4, T5, T6  
**Mar/18:** R1, R2  
**Mar/25:** R3, R4  
**Apr/1:** R5, R6  
**Apr/8:** A1, A2, A3, A4, A5, A6

**Topics Covered:** **Introduction, examples and structural properties** (2 weeks)

**Finite Horizon Markov Decision Processes** (2.5 weeks)

- Policies and optimality criteria
- Policy evaluation
- The principle of optimality
- Optimality of deterministic Markov policies
- Backward induction

**Infinite Horizon Markov Decision Processes** (2.5 weeks)

- Preliminary mathematics
- Value of a Policy
- Expected Total Discounted Reward
- Policy evaluation
- Optimality Equations
- Value Iteration
- Policy Iteration
- LP formulation and acceleration

**Simulation-based approach and reinforcement learning** (3 weeks)