An Introduction to Stochastic Control and Reinforcement Learning

structure of the course

Monday, July 7

08:30-09:00: Introduction to the course

09:00-10:30: PART 1 (Simone Garatti)
Finite state Markov chain (discrete-time) and Markov decision processes (MDP) (controlled Markov chain) and their applications. Discrete-time stochastic control. Finite horizon stochastic control problem, principle of optimality (Bellman equation)

11:00-12:30: PART 2 (Subhrakanti Dey)
Dynamic Programming and its solutions, Closed form solution for the Linear Quadratic Gaussian (LQG) control problem.

14:30-16:00: PART 3
Infinite horizon stochastic control problems (discounted and average cost with finite state and action space), Bellman optimality equation, existence of stationary control policy.

16:30-18:00: PART 4 (Simone Garatti)
Solution methodologies – value iteration and policy iteration and related algorithms.

Tuesday, July 8

09:00-10:30 + 11:00-12:30: PART 5 (S. Garatti)
Curse of dimensionality in solving Dynamic Programming algorithms, Approximate Dynamic Programming algorithms – approximation in policy space and value space, contraction properties and error bounds, simulation-based implementation.

14:30-16:00 + 16:30-18:00: PART 6 (S. Garatti)
Intro to reinforcement learning in the setting of MDP. Temporal difference methods (TD(0), TD(λ)), convergence properties. On-policy TD control (SARSA), Off-policy TD control such as Q-learning and its convergence properties, Applications (3 hrs with a break)

Wednesday, July 9

09:00-10:30: PART 7 (S. Garatti)
Advanced reinforcement learning. Value function approximation with Linear methods and function approximation. Deep reinforcement learning.

11:00-12:30: PART 8
policy gradient methods, actor-critic based reinforcement learning and their applications to continuous control (such as LQG) problems (3 hrs with a break)

REFERENCES

[1] D. P. Bertsekas, Dynamic Programming and Optimal Control, volumes 1 and 2, Athena Scientific, 2012

[2] D. P. Bertsekas, Reinforcement Learning and Optimal Control, Athena Scientific, 2019

[3] D. P. Bertsekas, A Course in Reinforcement Learning, 2nd edition, Athena Scientific, 2025 - available at: https://web.mit.edu/dimitrib/www/RLCOURSECOMPLETE%202ndEDITION.pdf

[4] D.P.Bertsekas and J.N. Tsitsiklis, Neuro-Dynamic Programming, Athena Scientific, 1996

[5] Sutton and Barto, Reinforcement Learning: Second Edition, MIT Press, 2018.

[6] Csaba Szepesvári, Algorithms for Reinforcement Learning, Morgan&Claypool publishers, 2010 – available at: https://sites.ualberta.ca/~szepesva/papers/RLAlgsInMDPs.pdf

For resources related to Bertsekas's lecture slides and videos, see also:
https://web.mit.edu/dimitrib/www/dpbook.html
https://web.mit.edu/dimitrib/www/RLbook.html

MATERIAL

Part 1: notes **"1. MDP_finite_horizon_control.pdf"**

Part 2: slides **"2-3. reinforcement_learning_bertinoro.pdf"**

Part 3: slides **"2-3. reinforcement_learning_bertinoro.pdf"**

Part 4: notes **"4. Value_Iteration_and_Policy_Iteration.pdf"**

Part 5: slides **"5. ADP_bertinoro.pdf"**

Part 6: notes **"6. Reinforcement_Learning_tab.pdf"**

Part 7: notes **"7. RL_with_function_approximation.pdf"**

Part 8: slides **"8. policy_gradient_combined_final.pdf"**
article https://arxiv.org/abs/1801.05039

All the material can be also downloaded from:

https://drive.google.com/drive/folders/1ypL-ZfEtNdhNSW-kI3F9IVJvHiv08NLi?usp=drive_link