Reinforcement Learning

Course Description

Reinforcement learning allows computers to derive problem-solving strategies without being explicitly programmed for the specific task, similar to the way humans and animals learn. After introducing the concepts of reinforcement learning, this course discusses the properties of Markov chains and single- and multi-armed bandits in detail. Special attention is given to the understanding of value functions and discounted value functions. The course connects reinforcement learning with neural networks and deep learning and discusses how Q-learning approaches can be used to utilize deep learning methods in reinforcement learning problems, including extensions such as double Q-learning, hierarchical learning, and actor–critic learning. Finally, the course discusses reinforcement learning approaches such as model-free and model-based learning and the tradeoff between exploration and exploitation.

Contents

1. Introduction to Reinforcement Learning
   1. Understanding Reinforcement Learning
   2. Components of Reinforcement Learning Systems
2. Markov Chains
   1. Markov Decision Process and Markov Property
   2. Value Functions and Discounted Value Functions
   3. General Utility Function
   4. Actions and Policy
   5. Bellman’s Equation
   6. Value Iteration
   7. Markov Chain Monte Carlo (MCMC)
3. Bandit
   1. Single-Armed Bandit
   2. Multi-Armed Bandit
4. Q-Learning
   1. Time-Difference Learning
   2. Reinforcement Learning with Neural Networks and Deep Q-Learning
   3. Experience Replay
   4. Double Q-Learning
   5. Delayed Sparse Rewards
   6. Hierarchical Learning
   7. Value- vs. Policy-Based Learning
   8. Actor–Critic Learning
5. Reinforcement Learning Approaches
   1. Model-Free Learning
   2. Model-Based Learning
   3. Exploration vs. Exploitation