Notes on reinforcement learning

Bart Frenk

December 17, 2020

Contents

1 Fundamentals of reinforcement learning

1.1 Week 1

1.1.1 Action-value methods for k-armed bandit problems

1.1.2 Incrementally computing action-values

- Use $Q_{n+1}(a) = Q_n(a) + \alpha_n(a)[R_n(a) - Q_n(a)]$

1.1.3 Non-stationary k-armed bandits

- Exponential recency-weighted average $\alpha_n(a) = \alpha$

1.1.4 Optimistic initial values

- To encourage exploration at the start

1.1.5 Upper-confidence-bound action selection

- Select action a that maximizes

$$Q_t(a) + c\sqrt{\frac{\ln t}{N_t(a)}}$$