Reinforcement Learning
Data Mining Project
732A65

Rasmus Holm

October 14, 2017

Abstract

Introduction

Theory

Sarsa

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t) \right]$$

Expected Sarsa

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \mathbb{E}\left[Q(S_{t+1}, A_{t+1})|S_{t+1} \right] - Q(S_t, A_t) \right]$$
$$= Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \sum_a \pi(a|S_{t+1}) Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t) \right]$$

Q-Learning

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_a Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t) \right]$$

Method

Result

Discussion

Conclusion