Information Processing and the Brain CW 2

Kheeran Naidu

December 2019

Introduction

For CW2, I have decided to implement and explore the behaviour of the classical *Reinforcement Learning* algorithm using *Temporal Difference Learning*. Reinforcement learning was developed from an area of psychology under the name trial-and-error learning [1] and the area of the optimal control problem in the form of Markov Decision Processes [**bellman1957dynamic**] [**bellman1957markov**]. Temporal-difference (TD) methods for reinforcement learning were proposed as a model of classical (or Pavlovian) conditioning in 1987 [2], and refined to the TD learning rule in 1990 [3].

Question 1

References

[1] Robert Sessions Woodworth. "Experimental Psychology. New York: Holt, 1938". In: *Department of Psychology Dartmouth College Hanover, New Hampshire* (1937).

[2] Richard S Sutton and Andrew G Barto. "A temporal-difference model of classical conditioning". In: *Proceedings of the ninth annual conference of the cognitive science society*. Seattle, WA. 1987, pp. 355–378.

[3] Richard S Sutton and Andrew G Barto. "Time-derivative models of pavlovian reinforcement." In: (1990).