Quiz 4

Nam Nguyen

October 2020

1 Introduction

The main objective of optimal policy is to determine the best action to take at each state, thus maximizing the rewards. Algorithms for this problem include two main iterations:

1. Function that determine the reward at each state, denoted $F(s)$

2. Initialize $F_0(s)$
   For $i = 1, 2, ..., H$, given $F_i(s)$,
   update $F_{(i+1)}(s) \leftarrow \max \sum T(s, a, s')(R(s, a, s') + F_i(s'))$

   The term in last equation is called Bellman update.

2 References

Advance Robotics Fall 2012, University of California at Berkeley.