Department of Computer Engineering

Artificial Intelligence

Mini Project 5 Theory Questions

Dr. Rohban

Parsa Mohammadian — 98102284

February 8, 2022

Contents

1 **1**
1.1 1
1.2 1
1.3 1
1.4 1

2 **1**
2.1 1
2.2 1
2.3 1

1

1.1

So the intuition is to show that from any starting point, value iteration with variable discount factor γ can be used to find the optimal policy. We know that $\inf -norm$ by definition is the largest (by size) value of a vector.

1.2

$$\gamma_k = 1 - \frac{1}{k+1} = \frac{k}{k+1}$$

$$\prod_{k=1}^{K} \gamma_k = \prod_{k=1}^{K} \frac{k}{k+1} = \frac{K!}{(K+1)!} = \frac{1}{K+1} \leq \frac{1}{K+1}$$

1.3

We already know that value iteration without discount factor ($\gamma = 1$) converges to an specific value. So it is obvious that value iteration with any discount factor near to one converges to an specific value. Here if we suppose value iteration with variable discount factor converges in K steps, then we can say that value iteration with constant discount factor γ_K converges to an specific value. And every other discount factor is smaller than γ_K. The smaller the discount factor, the faster we converge to an specific value. So if we replace the variable discount factor in our problem with γ_K the algorithm will converge. So we can be sure that the algorithm with variable discount factor will converge to an specific value too.

1.4

Variable discount factor can be useful in some of the problems. Actually it highly depends on how reward changes over time. If reward change over time is linear, I think constant discount factor is more suitable. Otherwise we can use variable discount factor corresponding to the reward change over time.

2

2.1

2.2

2.3