**Dependable Learning System Fall 2018 Homework 1**

**Chun-Hao Huang UIN:627002917**

This report presents the results of the Markov decision problem computing with three different methods including Value iteration, Gauss-Seidel with L2-norm convergence criterion and Dynamic programming with horizon H=4.

1. Value iteration

The number of iterations: 5

State value function V

|  |  |  |  |
| --- | --- | --- | --- |
| state | s1 | s2 | s3 |
| value | 0.4398 | 0.5125 | 0.4357 |

Deterministic policy

|  |  |  |  |
| --- | --- | --- | --- |
| state | s1 | s2 | s3 |
| action | a1 | a1 | a1 |

1. Gauss-Seidel with L2-norm convergence criterion

The number of iterations:5

State value function V

|  |  |  |  |
| --- | --- | --- | --- |
| state | s1 | s2 | s3 |
| value | 0.4505 | 0.5241 | 0.4483 |

Deterministic policy

|  |  |  |  |
| --- | --- | --- | --- |
| state | s1 | s2 | s3 |
| action | a1 | a1 | a1 |

1. Dynamic programming with horizon H=4

State value function V

|  |  |  |  |
| --- | --- | --- | --- |
| state | s1 | s2 | s3 |
| value | 0.4250 | 0.4977 | 0.4208 |

Deterministic policy

|  |  |  |  |
| --- | --- | --- | --- |
| state | s1 | s2 | s3 |
| action | a1 | a1 | a1 |

The results from these three methods are slightly different. The major difference between the result from value iteration and the result from finite horizon dynamic programming comes from the number of iteration. It takes five iterations for value iteration method to reach convergence. Compared to value iteration, finite horizon dynamic programming with H=4 takes only four stages to get its state value function. If we check their state function for every iteration, it would find that the state value function from finite horizon dynamic programming is the same as the fourth iteration result from value iteration. As for the difference between the result from value iteration and the result from value iteration by Gauss-Seidel, it comes from several aspects. In contrast to condition for value iteration to reach convergence, value iteration by Gauss-Seidel method reaches convergence when the result of L2-norm is less than epsilon. In addition, the state value function for value iteration function not only depends on Vk but also Vk+1. Therefore, the state value function from value iteration by Gauss-Seidel is greater than the state value function from value iteration.