

ECEN 689-605/700 Dependable Learning Systems Fall 2018

Homework 1

Due 1pm U.S. central time, September 10, Monday

Submission must be through ecampus

No late submission accepted!

A Markov Decision Process $S = \{s_1, s_2, s_3\}$, $A = \{a_1, a_2\}$. The state transition probability

matrices for a_1 and a_2 are $\begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.3 & 0.3 & 0.4 \\ 0.7 & 0.1 & 0.2 \end{bmatrix}$ and $\begin{bmatrix} 0.4 & 0.1 & 0.5 \\ 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$, respectively. The reward

functions under a_1 and a_2 are $\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.5 & 0.6 & -0.1 \\ 0.2 & 0.8 & 0.0 \end{bmatrix}$ and $\begin{bmatrix} 0.1 & 0.8 & 0.1 \\ -0.2 & 0.5 & 0.7 \\ 0.9 & -0.4 & 0.5 \end{bmatrix}$, respectively. Please

write software program in C/C++, Java or Python, to compute state value function and deterministic policy with discount factor $\gamma = 0.5$ and convergence criterion $\varepsilon = 0.02$. Please compute with the following three methods:

1. Value iteration (10 points)
2. Gauss-Seidel with L2-norm convergence criterion (10 points)
3. Dynamic programming with horizon $H = 4$ (10 points)

Please initialize value functions with all 0. Please submit your source code, and report the number of iterations and results. If the results from these three methods are different, please provide discussion (5 points).