Assignment 15

Miao-Chin Yen

March 8, 2022

Problem 1

Reference mrp_tdmc.py.
The result is given as follows:

- Tabular Monte-Carlo

$$\{A : 9.571428571428571, B : 5.642857142857143\}$$

- MRP
$$\{A : 12.93333333333333, B : 9.599999999999998\}$$

- Tabular TD(0)
$$\{A : 12.943540220953924, B : 9.59777231019758\}$$

- LSTD
$$\{A : 12.933333333333334, B : 9.600000000000001\}$$

They don't all give the same value function. MRP, Tabular TD(0) and LSTD give the similar value function and Tabular Monte-Carlo is significantly different from them. MRP should be the base line. Monte-Carlo just does the weighted sum. But in reality, doing average would not give enough information. Some value should be weighted more. Also, it uses trace experience. Therefore,the calculation process is restricted and may cause the value function to be not accurate enough.