P. Adam Aboud

500883647

CPS-824 Reinforcement Learning

**Assignment 2**

1. C. The accuracy was 100% for MC GLIE in the deterministic environment.

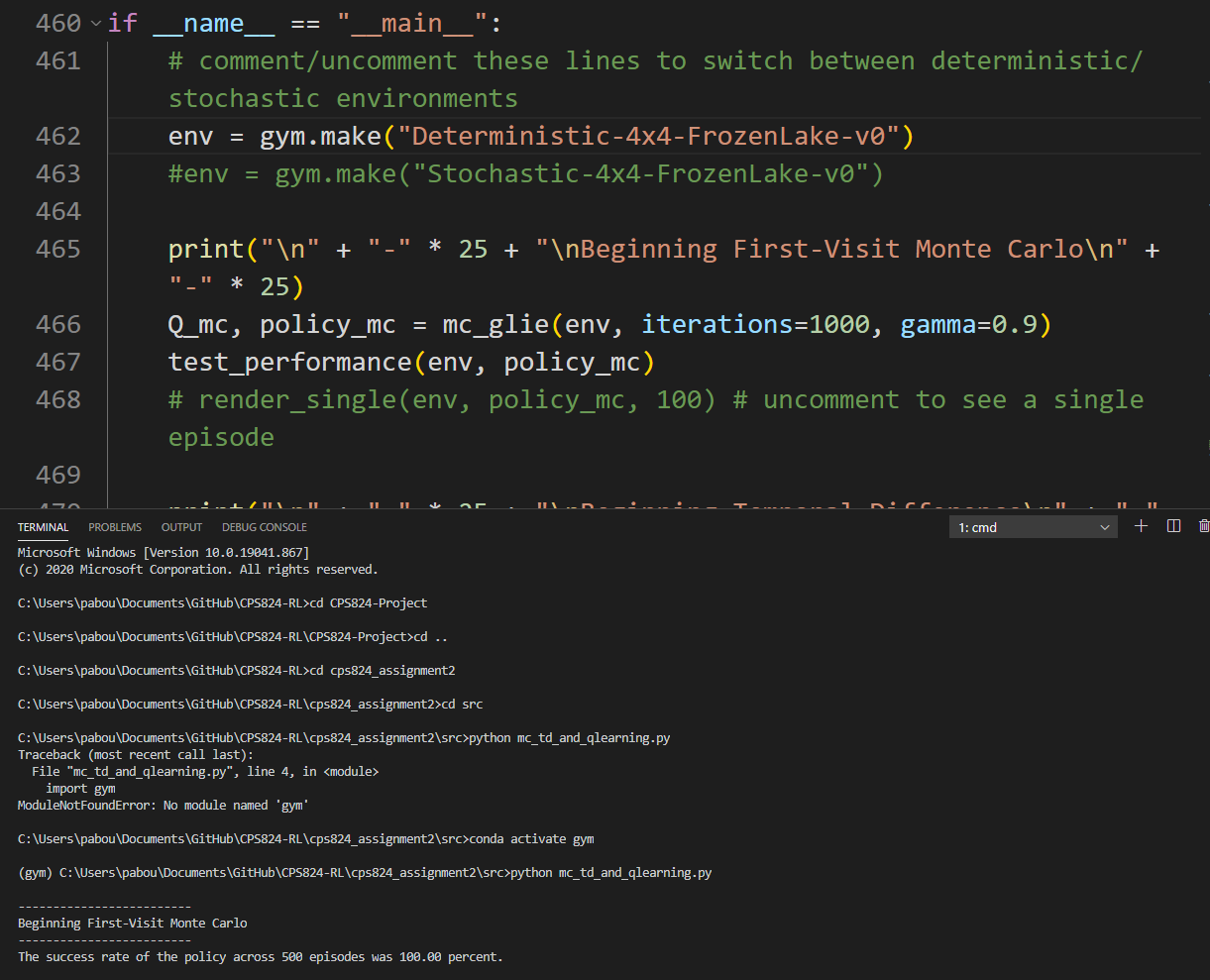


Fig. 1

1. D. The accuracy ranges anywhere from 5-30% at 1000 iterations. The reason the accuracy is so low is that the convergence for stochastic environments takes many more iterations. In order for the policy to be learned, it has to overcome the randomness inherent in the environment. This requires orders of magnitude greater iterations. Figure 2 shows MC GLIE with 1000 iterations, and the characteristic low accuracy rate. Figure 3 shows MC GLIE with 100 000 iterations.

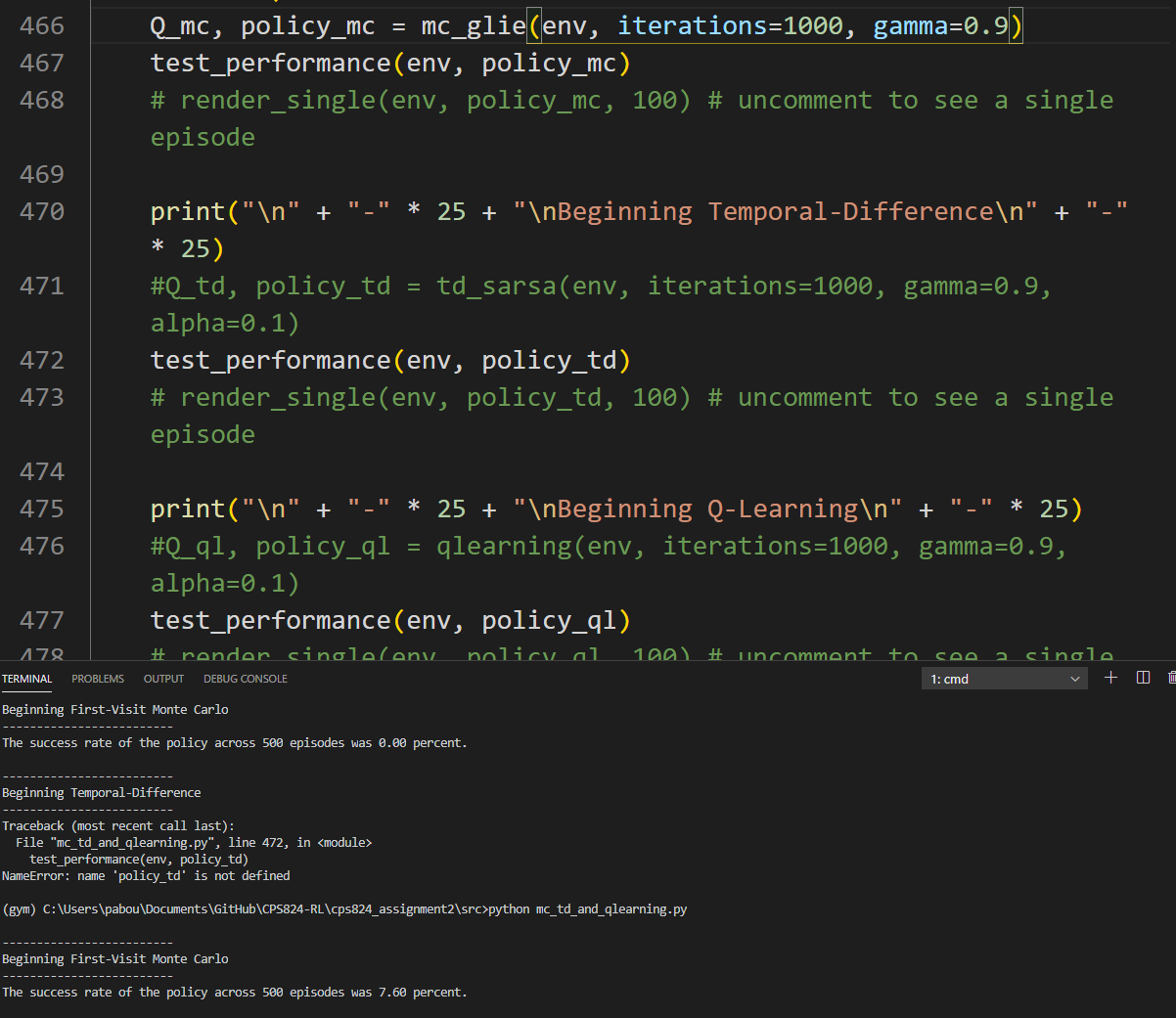


Figure 2.

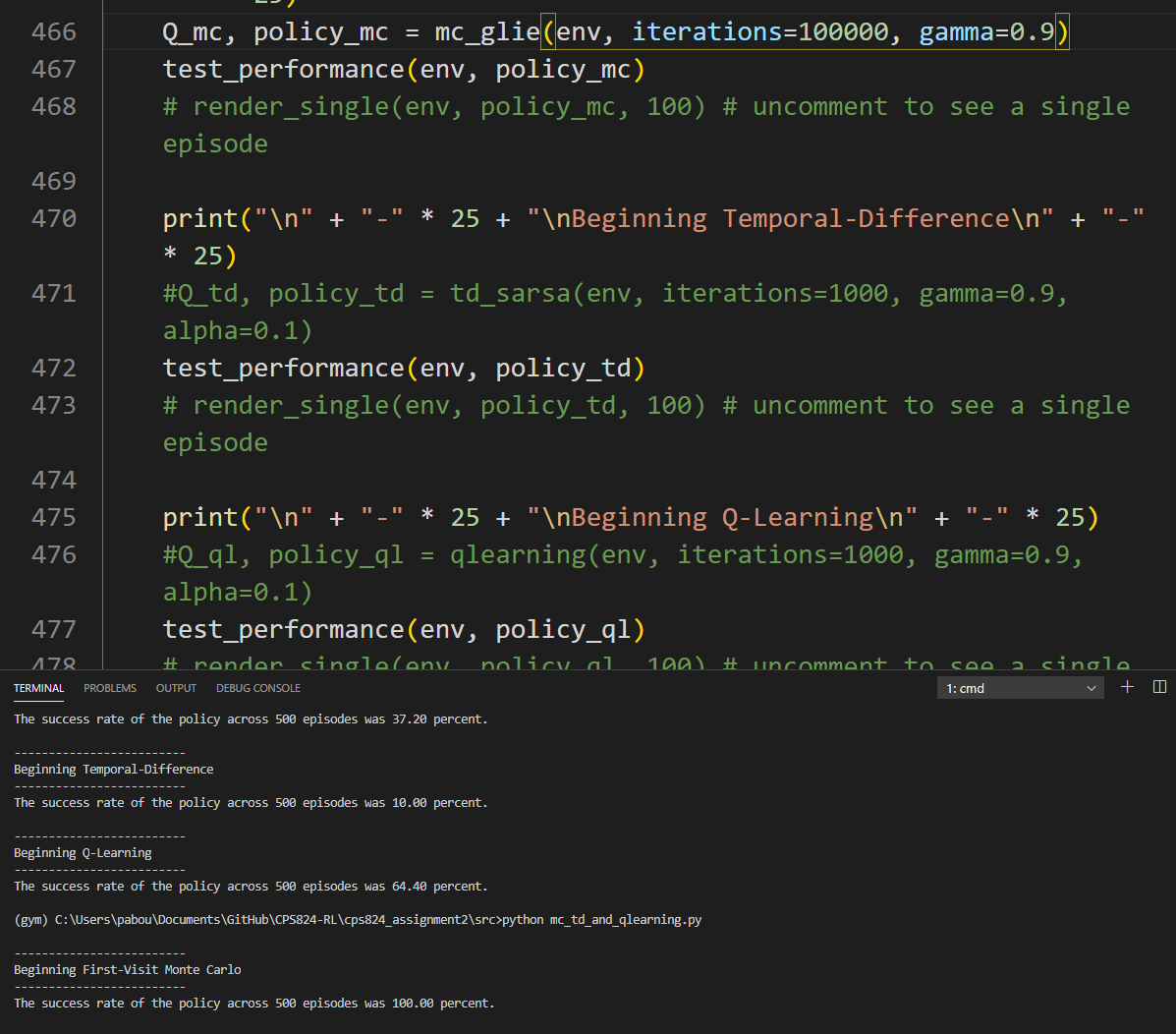
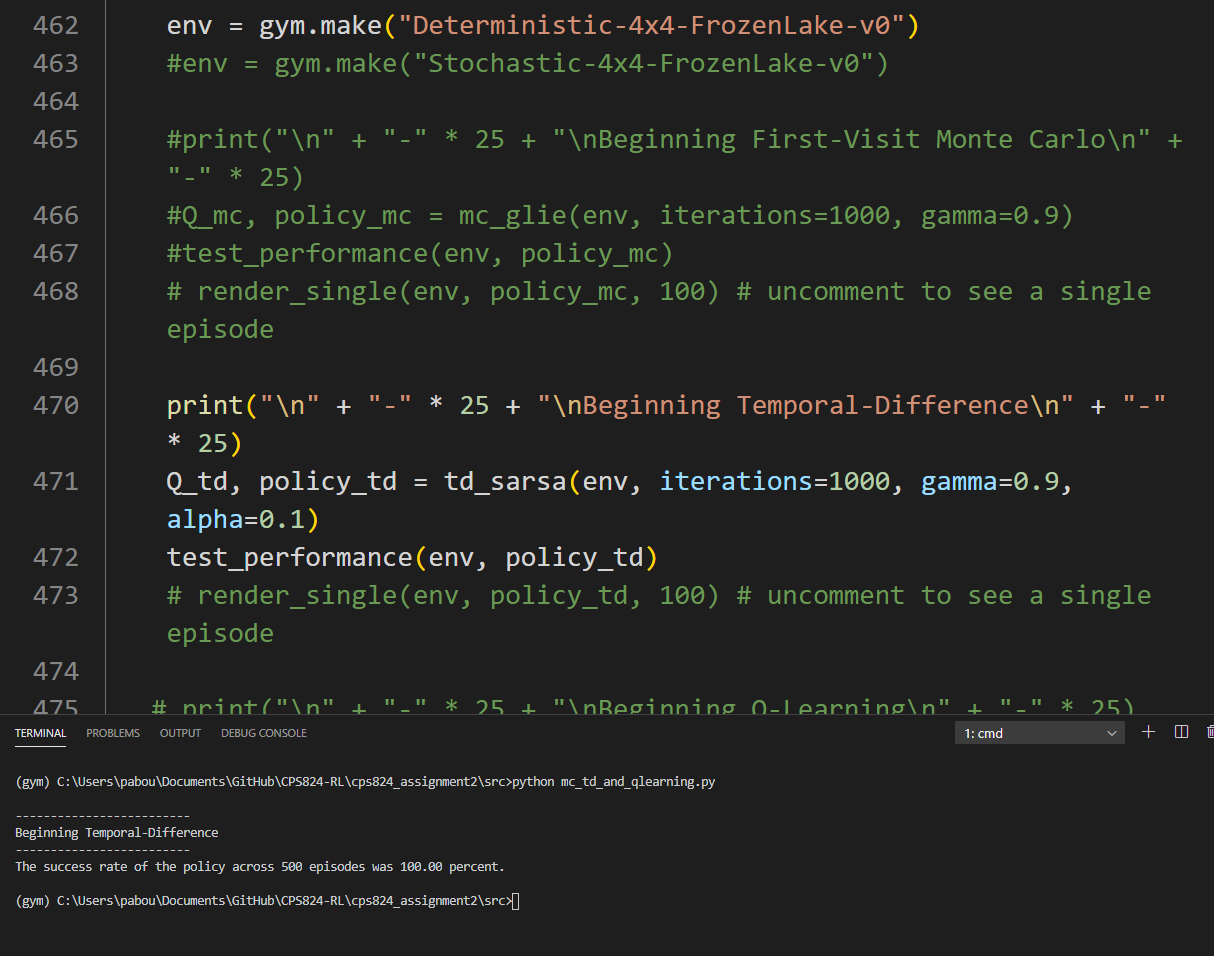
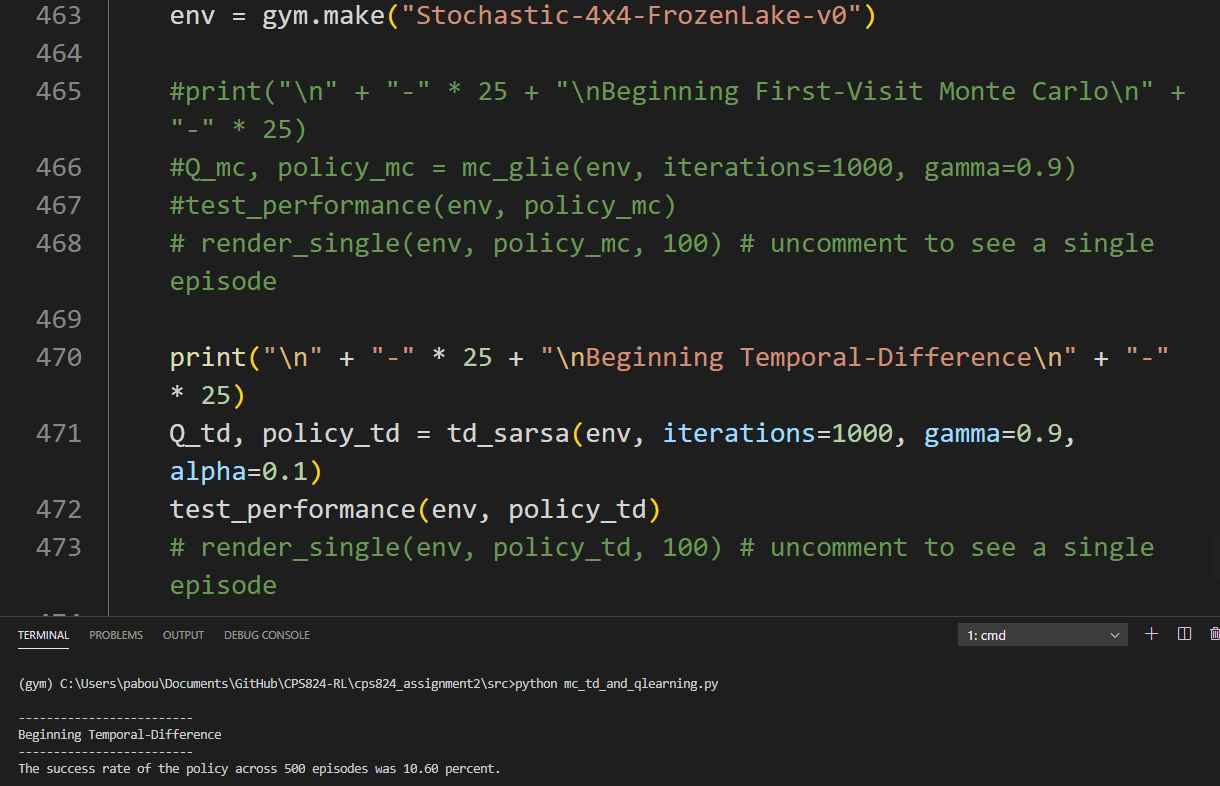
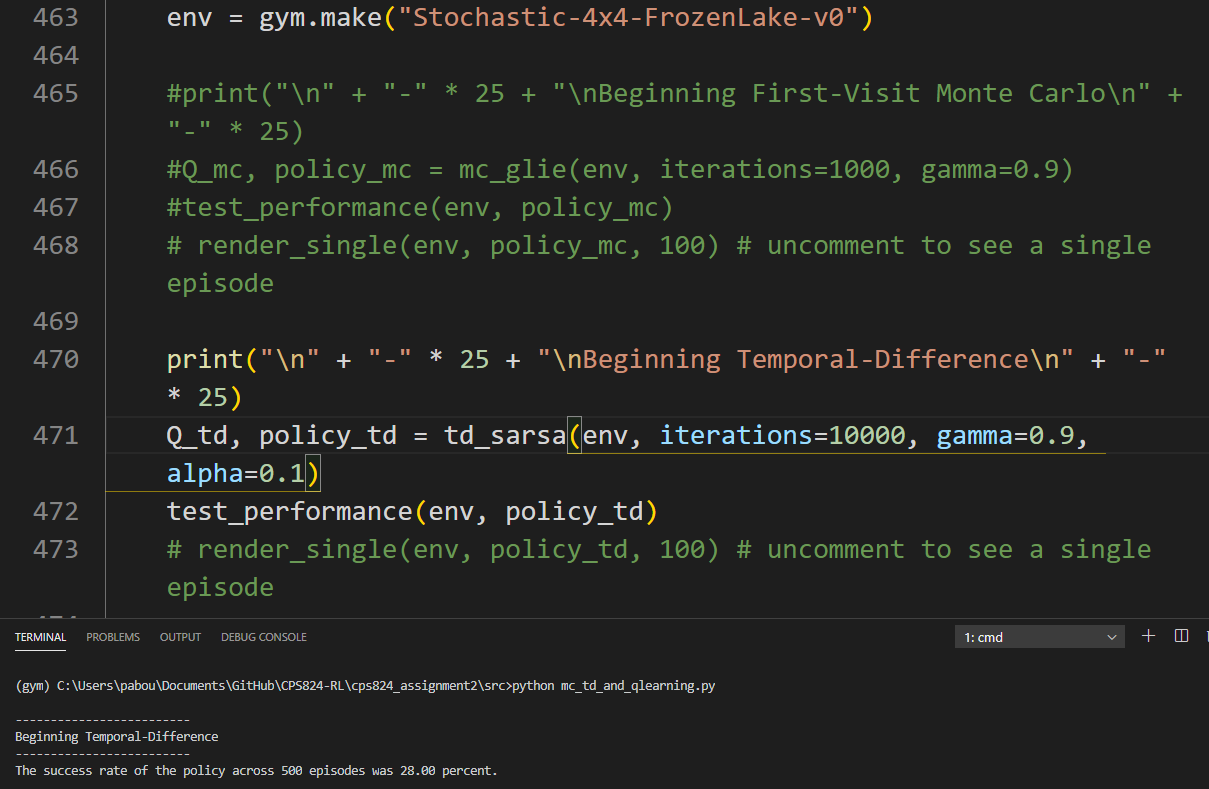


Figure 3.

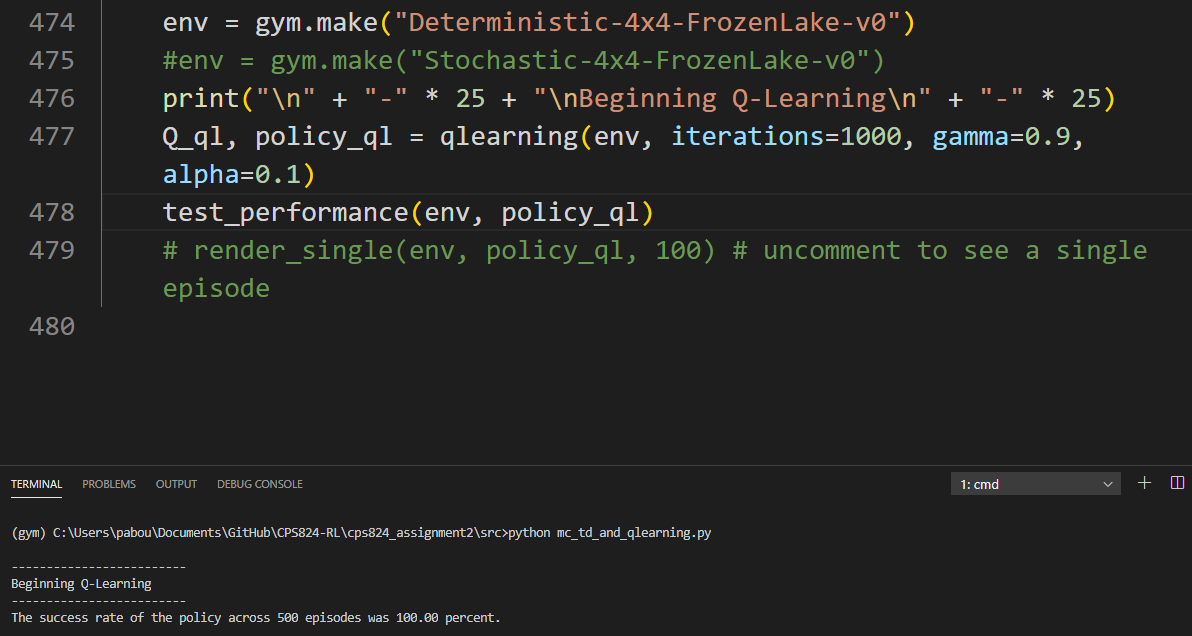
1. A. The accuracy for the deterministic environment using TD SARSA was 100%

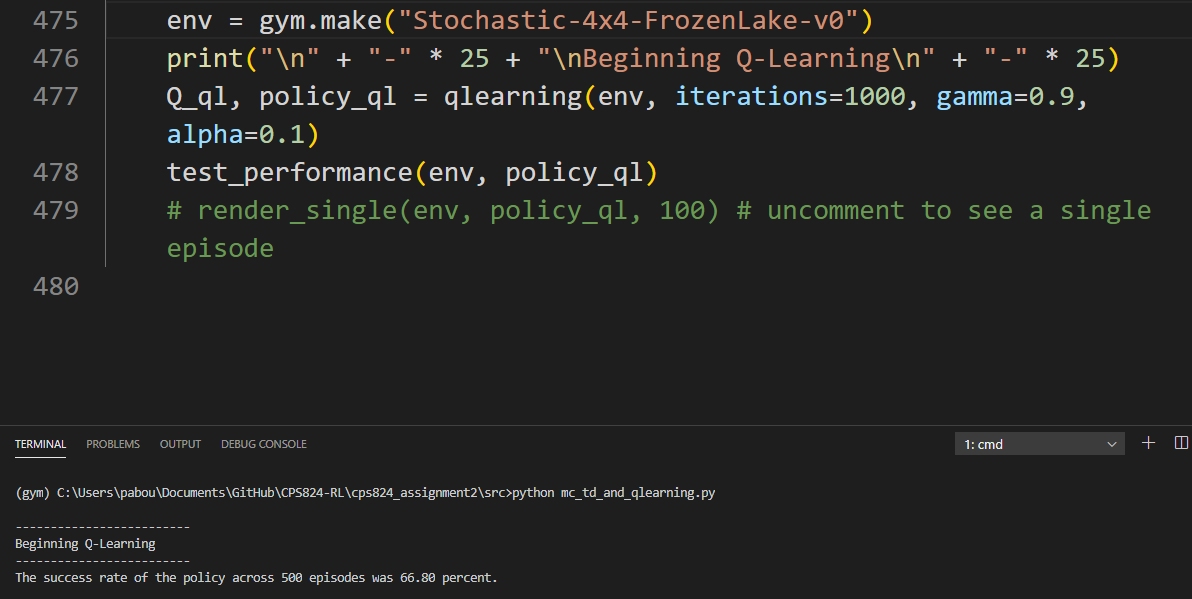






1. B. The accuracy for Q-Learning on the deterministic environment at 1000 iterations was 100%





Conclusions: