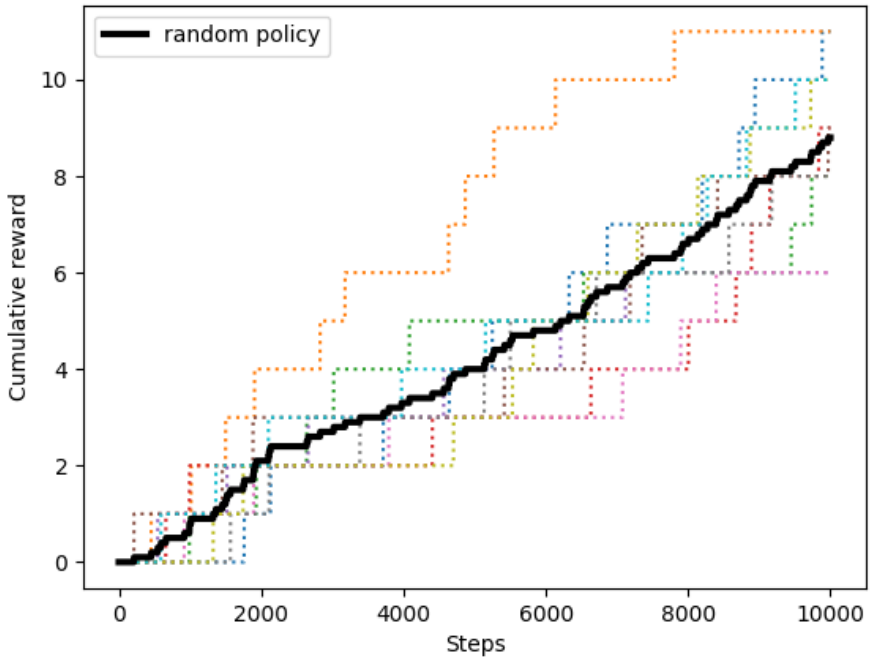
CS 5180: Exercise 0 Solutions

Gopi Sainath Mamindlapalli

Study Partner: Harin Kumar Nallaguntla

3. In a setting characterized by uncertainty, a strategy guided by human expertise is generally more effective than one based on random choices. When people make decisions, they can carefully weigh the options to select actions that are more likely to result in positive outcomes, thereby enhancing the likelihood of reaching the intended objective. On the other hand, a random approach lacks any form of strategic thinking, leading to unpredictable and often less-than-ideal results. Furthermore, human-driven strategies have the advantage of being able to evolve and adapt based on new information and experiences, while a random approach remains stagnant and unable to improve over time. This key distinction highlights the superior reliability of human-guided methods in achieving set goals when compared to random strategies.

Plot:



Cumulative reward vs Steps for Random policy

4. Better Policy:

The Strategy for this better policy is that it detects walls and corners and finds the path towards goal position. The policy becomes random policy when the agent goes to the middle of the room as there are no walls or corners around it but when it reaches any wall or corners then it follows the wall and tries to reach the goal position as quickly as possible.

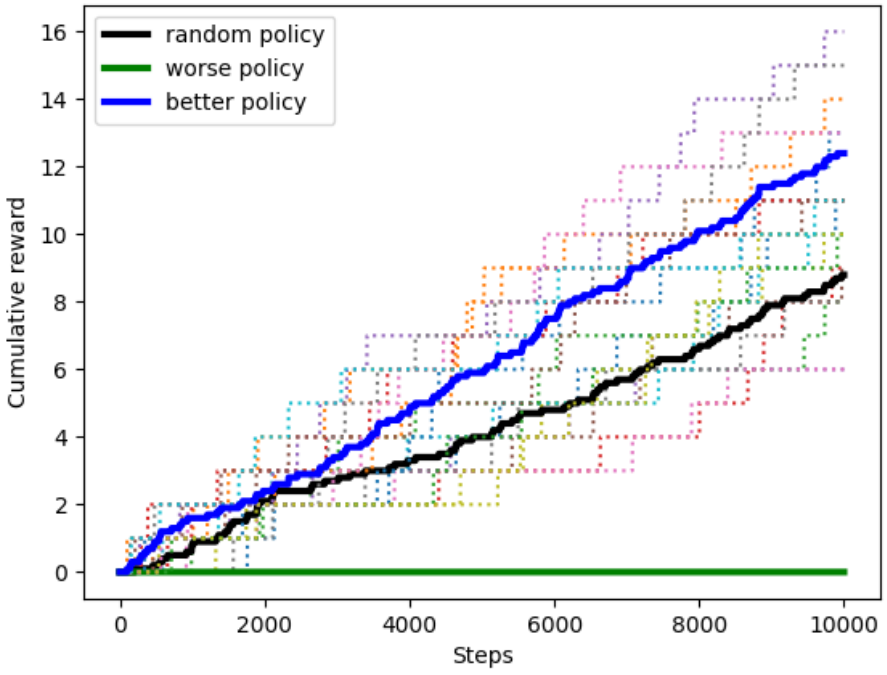
The performance of this policy is better than random policy because the probability of taking the correct action in random policy is 0.25, whereas for this better policy the probability of taking the correct action becomes 0.5 when the agent reaches the walls or corners. In this way the performance is increased.

Worse policy:

The strategy for the worse policy is simple as it always goes to right side regardless of the present state of the agent. Reaching goal position and getting a reward is almost impossible for this policy as it doesn't adapt according to the present state of the agent.

The performance of this policy is worse than random policy because after sometime of the iterations this policy always takes the wrong steps which doesn't take the agent to the goal position.

Plot:



Cumulative reward vs Steps for Better policy, Random policy and Worse policy