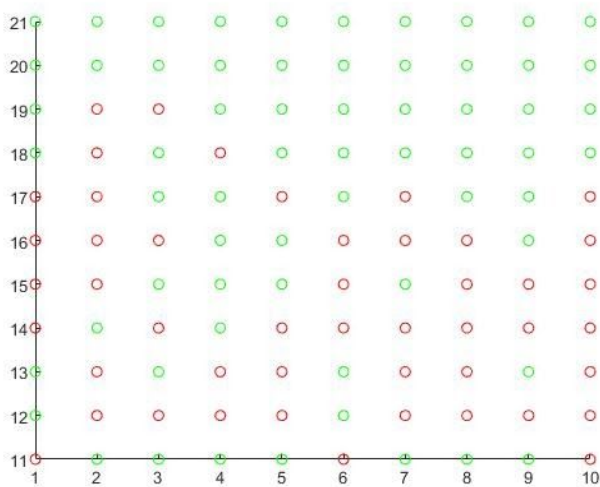**Reinforcement learning: assignment 4**
Vishal I B, 917809310

The policy was created as a cell array of lists. With every list having probabilities of the two actions. For the player sum from 11 to 21, the policy was hit and for 20 and 21, the policy was stick. The Q cell array was created and initialized with zeros. Returns was a cell array with a list for every action. For every episode, the states were generated using the current policy. And the rewards were determined at the end. After the generation of the states, the episode was looped from last time step. G was calculated with the rewards and then appended to returns cell array for that particular action. Q for every state was the average of returns from the action list of the return cell array. The max of the two action was made one in the policy.

Useable ace and no hit for soft 17:

graph:

policy:

[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[1,0]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[0,1]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]
[0,1]	[1,0]	[1,0]	[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[1,0]
[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[0,1]	[1,0]
[1,0]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]	[1,0]	[0,1]	[1,0]
[1,0]	[1,0]	[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]
[0,1]	[1,0]	[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]

No useable ace,no hit on soft-17

[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[1,0]	[1,0]	[0,1]	[1,0]
[0,1]	[1,0]	[1,0]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]	[0,1]
[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[0,1]	[0,1]	[1,0]
[0,1]	[0,1]	[1,0]	[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]
[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]
[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]

No useable ace, hit on soft-17

[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[1,0]	[1,0]	[0,1]	[1,0]
[0,1]	[1,0]	[1,0]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]	[0,1]
[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[0,1]	[0,1]	[1,0]
[0,1]	[0,1]	[1,0]	[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]
[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]
[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]

Useable ace, hit on soft-17

[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[1,0]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[0,1]
[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]
[0,1]	[1,0]	[1,0]	[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[1,0]
[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[0,1]	[1,0]	[1,0]	[0,1]	[1,0]
[1,0]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[1,0]	[1,0]
[1,0]	[1,0]	[1,0]	[0,1]	[0,1]	[1,0]	[1,0]	[1,0]	[0,1]	[1,0]
[1,0]	[1,0]	[0,1]	[0,1]	[1,0]	[0,1]	[1,0]	[0,1]	[0,1]	[1,0]
[0,1]	[1,0]	[0,1]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[1,0]	[1,0]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]