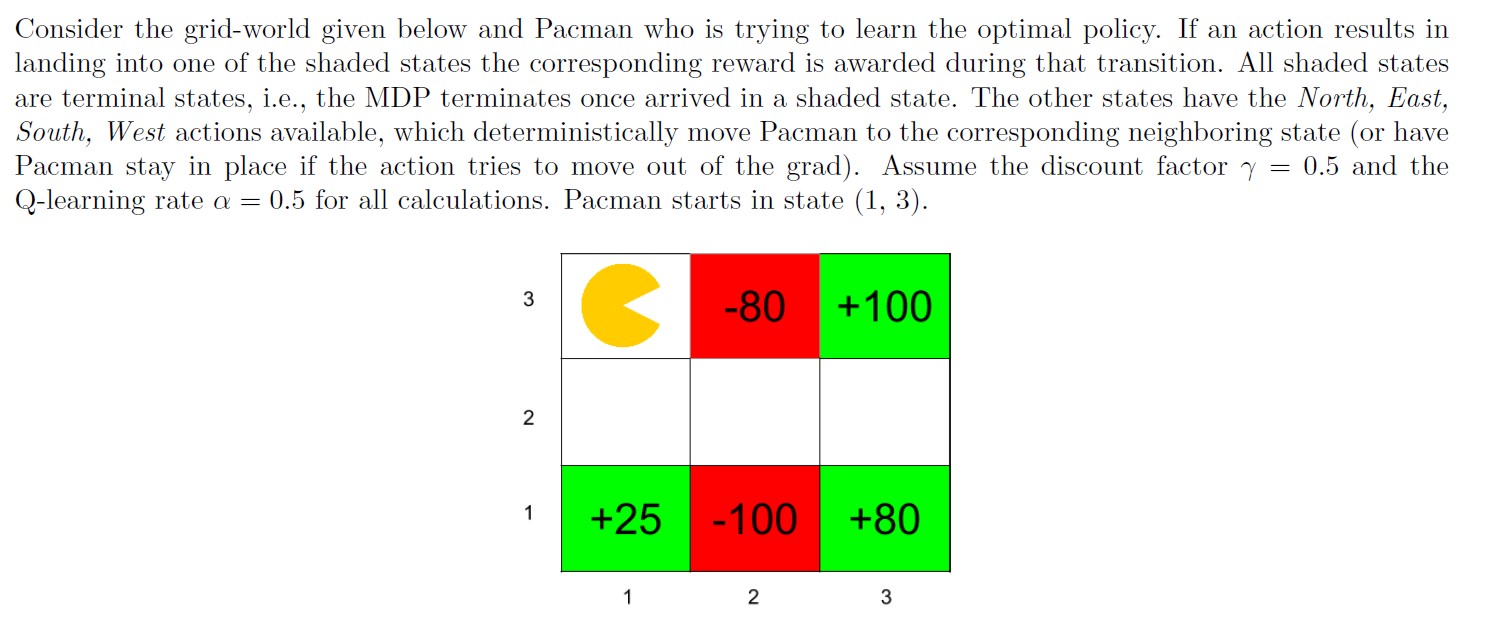
**CAP 6629 Homework 1**

**Due: Feb-11, 2022 (11:59PM)**

Hint:



1. What is the value of optimal value function V\* at the following states:

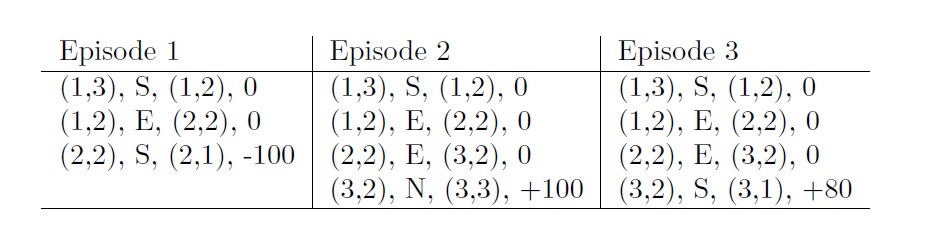
V\*(3,2) = **100 (only one step to reach maximum)**

V\*(2,2)= **50 = 0+ γ0.5 \*0 step 1-right 🡪 0+ γ0.5 \*100 🡪 step 2-up = 50**

V\*(1,3)= **12.5 0+ γ0.5 \*0 step 1-up 🡪 [ +25 or +100 same] Max = 0.5 \*25 =12.5 || 0.5\*100 = 50 \*0.5 =25\*0.5 =12.5 [1,1]**

Shaun P note: I would like clarification on this last V\*(1,3)

1. The agent starts from the top left corner and you are given the following episodes from runs of the agent through this grid-world. Each line is an Episode is a tuple containing *(s,a,s’,r).* Hint: N, E, S, W refer to the moving directions.



Using Q-learning updates, what the following Q-values after the above three episodes: Q((3,2),N)=  **50**

Q((1,2),S)= **0**

Q((2,2),E)= **12.5**