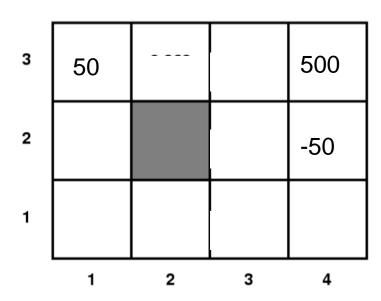# CSCI 561
Foundation for Artificial Intelligence

Discussion Section
(Week 4)

PROF WEI-MIN SHEN WMSHEN@USC.EDU

MDP

Given a Gridworld domain, where terminal states (1,3), (4,3), and (4,2) have rewards 50, 500, and -50 respectively, the set of possible actions are {N,E,S,W, or X for terminal states}, the agent moves deterministically, all V and Q values for non-terminal states have been initialized to 0.0, answer the questions below.

3	50	~ ~~~		500
2				-50
1				
	1	2	3	4

Circle the letter that corresponds to the best answer for the question.

What are the optimal utilitu values V for each state in the above grid if $\gamma = 0.5$, c(a)=0, R(s)=0 for the non terminal states?

(Remember $V_{t+1}(s) = R(s) + Max_{a \varepsilon A}\{c(a) + \gamma \Sigma_{s' \varepsilon S} P(s'|a,s) V_t(s')\}$)

a. $V_{(1,1)}$=15.75, $V_{(1,2)}$=25, $V_{(2,1)}$=31.25, $V_{(2,3)}$=125, $V_{(3,1)}$=62.5, $V_{(3,2)}$=125, $V_{(3,3)}$=250, $V_{(4,1)}$=25
b. $V_{(1,1)}$=12.5, $V_{(1,2)}$=25, $V_{(2,1)}$=31.25, $V_{(2,3)}$=125, $V_{(3,1)}$=62.5, $V_{(3,2)}$=125, $V_{(3,3)}$=250, $V_{(4,1)}$=31.25
c. $V_{(1,1)}$=15.625, $V_{(1,2)}$=25, $V_{(2,1)}$=31.25, $V_{(2,3)}$=125, $V_{(3,1)}$=62.5, $V_{(3,2)}$=125, $V_{(3,3)}$=250, $V_{(4,1)}$=31.25
d. $V_{(1,1)}$=12.5, $V_{(1,2)}$=25, $V_{(2,1)}$=25, $V_{(2,3)}$=25, $V_{(3,1)}$=50, $V_{(3,2)}$=100, $V_{(3,3)}$=250, $V_{(4,1)}$=25
e. None of the above

C

Q-Learning

What are the Q values of state (3,2) in the above grid if $\gamma = 0.5$, c(a)=0, R(s)=-2 for non-terminal states?

(Remember $Q_{t+1}(a,s) = R(s) + c(a) + \gamma \Sigma_{s' \varepsilon S} P(s'|a,s) \max_{a' \varepsilon A} Q_t(a's')$)

a. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=59
b. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=27.5
c. $Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S)}$=62.5
d. $Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=31.5
e. **None of the above**

50			500
	(grey)		-50

3   2   1

1   2   3   4

Q-Learning

What are the Q values of state (3,2) in the above grid if $\gamma = 0.5$, c(a)=0, R(s)=-2 for non terminal states?

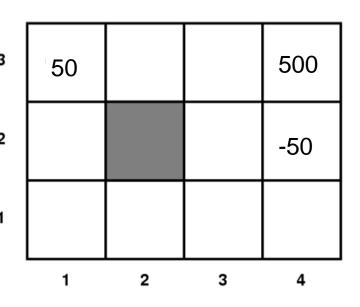(Remember $Q_{t+1}(a,s) = R(s) + c(a) + \gamma \Sigma_{s' \varepsilon S} P(s'|a,s) \max_{a' \varepsilon A} Q_t(a's')$)

a. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=59
b. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=27.5
c. $Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S)}$=62.5
d. $Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=31.5
e. None of the above

50		E248	500
			-50

Q-Learning

What are the Q values of state (3,2) in the above grid if γ = 0.5, c(a)=0, R(s)=-2 for non terminal states?

(Remember $Q_{t+1}(a,s) = R(s) + c(a)+\gamma\Sigma_{s'\epsilon S} P(s'|a,s)\max_{a'\epsilon A} Q_t(a's')$)

a. $Q_{((3,2),N)}=122$, $Q_{((3,2),E)}=-27$, $Q_{((3,2),S)}=59$
b. $Q_{((3,2),N)}=122$, $Q_{((3,2),E)}=-27$, $Q_{((3,2),S)}=27.5$
c. $Q_{((3,2),N)}=125$, $Q_{((3,2),E)}=-25$, $Q_{((3,2),S)}=62.5$
d. $Q_{((3,2),N)}=120$, $Q_{((3,2),E)}=-27$, $Q_{((3,2),S)}=31.5$
e. None of the above

	1	2	3	4
3	50		E248	500
2		(grey)	N122	-50
1				

Q-Learning

What are the Q values of state (3,2) in the above grid if γ = 0.5, c(a)=0, R(s)=-2 for non terminal states?

(Remember $Q_{t+1}(a,s) = R(s) + c(a)+\gamma\Sigma_{s'\varepsilon S} P(s'|a,s)\max_{a'\varepsilon A} Q_t(a's')$)

a. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=59
b. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=27.5
c. $Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S)}$=62.5
d. $Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=31.5
e. None of the above

B

50		E248	500
		N122 S27.5	-50
		N59	