

# CSE 5525 Artificial Intelligence II

## Quiz #2: Markov Decision Process

Wei Xu, Ohio State University

Your Name: \_\_\_\_\_ OSU Username: \_\_\_\_\_

### 1 Markov Decision Processes

#### Questions:

- 1) Write out the equations to be used to compute  $Q_i^*$  from  $R, T, V_{i-1}^*, \gamma$  and to compute  $V_i^*$  from  $R, T, Q_i^*, \gamma$ .

$$Q_i^*(s, a) = \sum_{s'} T(s, a, s') (R(s, a, s') + \gamma V_{i-1}^*(s'))$$

$$V_i^*(s) = \max_a Q_i^*(s, a)$$

- 2) Consider the MDP with transition model and reward function as given in the table below. Assume the discount factor  $\gamma = 1$ , i.e., no discounting. Fill in the values for  $V_0^*, V_1^*, V_2^*, Q_1^*, Q_2^*$  in the graph below.

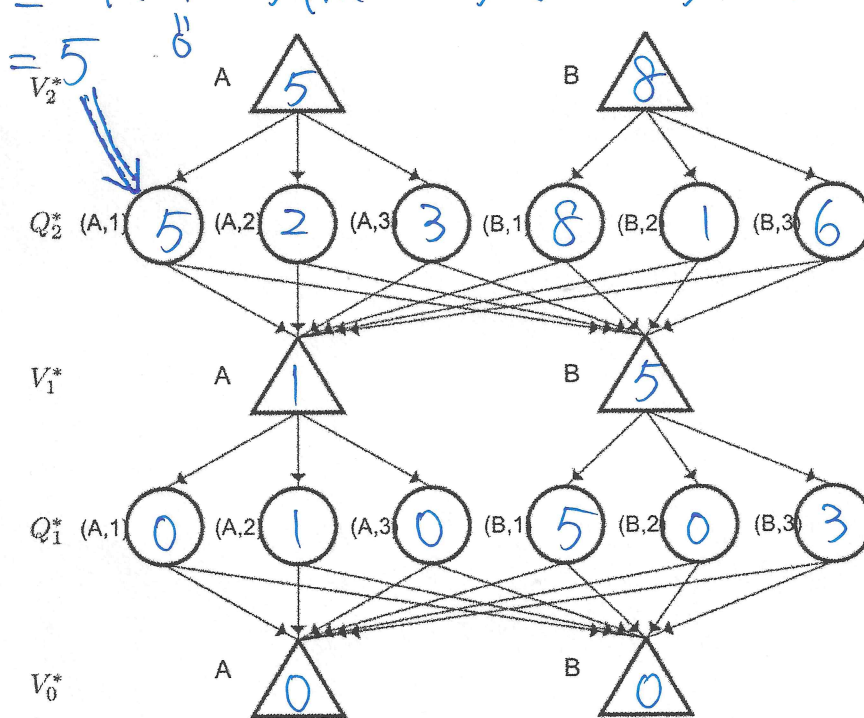
s	a	s'	T(s,a,s')	R(s,a,s')
A	1	A	0	0
A	1	B	1	0
A	2	A	1	1
A	2	B	0	0
A	3	A	0.5	0
A	3	B	0.5	0

s	a	s'	T(s,a,s')	R(s,a,s')
B	1	A	0.5	10
B	1	B	0.5	0
B	2	A	1	0
B	2	B	0	0
B	3	A	0.5	2
B	3	B	0.5	4

Here is an example:

$$Q_2^*(A, 1) = \sum_{s' \in \{A, B\}} T_{s'}(R + \gamma V_1^*(s'))$$

$$= T(A, 1, A) (R(A, 1, A) + \gamma V_1^*(A)) + T(A, 1, B) (R(A, 1, B) + \gamma V_1^*(B))$$



3) Let  $\pi_i^*(s)$  be the optimal action in state  $s$  with  $i$  time steps to go. Fill in the following tables:

$s$	$\pi_1^*(s)$
A	2
B	1

$s$	$\pi_2^*(s)$
A	1
B	1