

**Question 4. Bayes Nets**

A professor is heading back into the lab late at night and is trying to guess whether his student is working on their paper. He has noticed that whenever this student is working, their car is in the garage and there is music playing in the lab. Let  $W$ ,  $C$ ,  $M$  be three binary random variables corresponding to the events “student Working on paper”, “Car in parking lot”, and “Music playing”. Through careful observation, the professor has determined the following conditional probabilities:

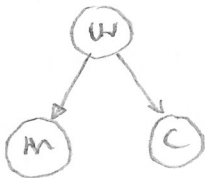
$$(1) \quad P(C|W, M) = P(C|W) = \begin{array}{c|cc} & W=1 & W=0 \\ \hline C=1 & 0.8 & 0.1 \\ C=0 & 0.2 & 0.9 \end{array}$$

$$(2) \quad P(M|W) = P(C|W)$$

$$(3) \quad P(W=1) = 0.8$$

Note that equation (2) says that  $C$  and  $M$  have the same conditional distribution (e.g.  $P(C=1|W=1) = P(M=1|W=1) = 0.8$ .)

(a) Draw a Bayesian network model (i.e. a directed graphical model) for this problem.



(b) Compute the marginal distributions  $P(C)$  and  $P(M)$ , before any evidence is available

$$P(C, W) = P(C|W)P(W) = P(M|W)P(W) = P(M, W) \quad \text{so joint pdf's also same}$$

$$\begin{array}{l} \xrightarrow{C} \begin{array}{cc} W \\ \left[ \begin{array}{cc} 0.64 & 0.02 \\ 0.16 & 0.18 \end{array} \right] \end{array} \rightarrow \begin{array}{c} \left[ \begin{array}{c} 0.64 \\ 0.34 \end{array} \right] \\ C=1 \\ C=0 \end{array} \\ \uparrow P(C) \text{ or } P(M) \end{array}$$

(c) Upon entering the garage, the professor notices that the student's car is parked there. Calculate  $P(W|C=1)$ .

Bayes Rule:

$$P(W|C=1) = \frac{P(C=1|W)P(W)}{P(C=1)} = \frac{1}{0.66} \times \begin{bmatrix} 0.8 \times 0.8 \\ 0.2 \times 0.1 \end{bmatrix} = \begin{bmatrix} 0.97 \\ 0.03 \end{bmatrix} \begin{matrix} W=1 \\ W=0 \end{matrix}$$

$\uparrow$   $\uparrow$   $\uparrow$   
 $P(C=1)$   $P(W)$   $P(C=1|W)$   
 (from marginal)

(d) Calculate the updated probability that music is playing in the lab

We need to take  $C=1$  into account

$$P(M, W|C=1) = P(M|W, C=1)P(W|C=1) = P(M|W)P(W|C=1)$$

$$\begin{aligned}
 &= \begin{matrix} & W \\ C & \begin{matrix} 1 & 0 \end{matrix} \\ \begin{matrix} 1 \\ 0 \end{matrix} & \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \end{matrix} \times \begin{matrix} W|C=1 \\ \begin{bmatrix} 0.97 \\ 0.03 \end{bmatrix} \end{matrix} \\
 &= \begin{bmatrix} 0.8 \times 0.97 & 0.1 \times 0.03 \\ 0.2 \times 0.97 & 0.9 \times 0.03 \end{bmatrix} = \begin{matrix} M \backslash W \\ \begin{matrix} 1 & 0 \end{matrix} \\ \begin{matrix} 1 \\ 0 \end{matrix} & \begin{bmatrix} 0.78 & 0.003 \\ 0.19 & 0.027 \end{bmatrix} \end{matrix} \rightarrow \begin{matrix} P(M|C=1) \\ \begin{bmatrix} 0.783 \\ 0.217 \end{bmatrix} \end{matrix}
 \end{aligned}$$

**Question 5. Decision Tree Learning**

Consider the following dataset for training a decision tree:

	$A_1$	$A_2$	$A_3$	$A_4$	Output
1	1	1	1	0	1
2	1	0	1	0	1
3	0	0	1	1	1
4	0	0	0	0	1
5	1	0	0	1	0
6	1	1	1	1	0
7	1	0	0	0	0
8	0	0	1	1	0

$$H = B(1/2) = 1$$



change to 1

(a) The attribute  $A_4$  is selected for the first (root) node of the decision tree. Calculate the information gain based on splitting on  $A_4$ .

$$H(O|A_4) = \frac{1}{2} B(1/4) + \frac{1}{2} B(3/4) = 0.8113$$

$$\text{Gain} = 1 - 0.8113 = 0.1887$$

(b) Consider the **TRUE** branch of the  $A_4$  split. Identify whether an additional split is needed, and if so which attribute to select next.

$$A_1: \frac{1}{2} B(0) + \frac{1}{2} B(\frac{1}{2}) = 0.5 \quad \rightarrow \text{CHOOSE } A_1$$

$$A_2: \frac{1}{4} B(0) + \frac{3}{4} B(\frac{1}{3}) = 0.6887$$

$$A_3: \frac{3}{4} B(\frac{1}{3}) + \frac{1}{4} B(0) = 0.6887$$

c) FALSE BRANCH

$$A_1: \frac{3}{4} B(\frac{2}{3}) + \frac{1}{4} B(1) = 0.6887$$

$$A_2: \frac{1}{4} B(1) + \frac{3}{4} B(\frac{2}{3}) = 0.6887$$

$$A_3: \frac{1}{2} B(1) + \frac{1}{2} B(\frac{1}{2}) = 0.5 \quad \rightarrow \text{choose } A_3$$

$$d) A_3: \frac{5}{8} B(\frac{3}{5}) + \frac{3}{8} B(\frac{1}{5})$$

$$\rightarrow \frac{3}{4} B(\frac{2}{3}) + \frac{1}{4} B(0) = 0.6887 \quad \checkmark$$

**Question 6. Markov Decision Processes**

Consider the grid world shown below. The agent starts in state  $S_1$  and is trying to reach the positive goal  $G_2$ . There is also a negative terminal  $G_1$  which the robot wants to avoid. The reward for each of the two goals and the four states is shown in the top right of each square. Assume that there are *three* action choices in each state: *Left*, *Up*, and *Down*. Use the standard noise model for actions in which the probability of moving in the desired direction is 0.8 and there is 0.1 chance to move in each of the orthogonal directions (this is the same action model we used in class and is used in your book). Assume that rewards are not being discounted (i.e. discount factor of 1.0).

+1.0 $G_2$	-0.01 $S_4$ ←	-0.01 $S_3$ ←	
	-1.0 $G_1$	-0.01 $S_2$ ↑	-0.01 $S_1$ ↓

(a) Assume that the utilities for each of the states  $S_1$  to  $S_4$  are initialized to 0.1. Perform two rounds of value iteration, and show the utilities for each of the four states at each round.

Round 1:

$$U(S_1) = -0.01 + \max_{LUD} \{ 0.8 \times (0.1) + 0.1 \times (0.1) + 0.1 \times (0.1) \} \quad \begin{matrix} \leftarrow \\ \uparrow \\ \downarrow \end{matrix} \quad \begin{matrix} 0.1 \\ 0.09 \end{matrix}$$

All other utilities will be same by symmetry  $U \neq D$

$$U(S_2) = -0.01 + \max_{LUD} \{ 0.8 \times (-1) + 0.2 \times (0.1) \} \quad \begin{matrix} \leftarrow \\ \uparrow \\ \downarrow \end{matrix} \quad \begin{matrix} -0.78 \\ -0.01 \end{matrix}$$

$0.8 \times (0.1) + 0.1 \times (-1) + 0.1 \times (0.1)$  (U) TIED w/ D  
D will be same by symmetry

$$U(S_3) = -0.01 + \max_{LUD} \{ \text{ALL OPTIONS SAME, IDENTICAL TO } S_1 \} \quad \begin{matrix} \leftarrow \\ \uparrow \\ \downarrow \end{matrix} \quad \begin{matrix} 0.09 \\ 0.09 \end{matrix}$$

(L) TIED w/ U & D

$$U(S_4) = -0.01 + \max_{LUD} \{ 0.8 \times (1.0) + 0.1 \times (0.1) + 0.1 \times (-1) \} \quad \begin{matrix} \leftarrow \\ \uparrow \\ \downarrow \end{matrix} \quad \begin{matrix} 0.71 \\ 0.7 \end{matrix}$$

$0.9 \times (0.1) + 0.1 \times (1.0)$  (U)  $0.19$   
 $0.8 \times (-1) + 0.1 \times (1) + 0.1 \times (0.1)$  (D)  $-0.69$

Round 2:

$$U(S_1) = -0.01 + \max_{LUD} \{ \begin{array}{l} \text{L IS NOT COMPETITIVE W/ NEGATIVE UTIL} \\ \text{L} \end{array} \}$$

$\uparrow \quad \rightarrow \quad \leftarrow \quad \rightarrow$   
 $0.9 \times (0.09) + 0.1 \times (-0.02) \quad U \quad (0.069) \quad D \quad \text{BREAK TIE ALPHABET}$

D will be same by symmetry

$$U(S_2) = -0.01 + \max_{LUD} \{ \begin{array}{l} \text{L CAN'T WIN} \\ \text{L} \end{array} \}$$

$\uparrow \quad \leftarrow \quad \rightarrow$   
 $0.8 \times (0.09) + 0.1 \times (-1) + 0.1 \times (0.09) \quad (U) \quad (-0.029)$

D CAN'T WIN BECAUSE  $0.09 > -0.02$

$$U(S_3) = -0.01 + \max_{LUD} \{ \begin{array}{l} \text{L CAN'T WIN} \\ \text{L} \end{array} \}$$

$\leftarrow \quad \downarrow \quad \uparrow$   
 $0.8 \times (0.7) + 0.1 \times (-0.02) + 0.1 \times (0.09) \quad (L) \quad (0.557)$

U & D CAN'T WIN

$$U(S_4) = -0.01 + \max_{LUD} \{ \begin{array}{l} \text{L CAN'T WIN} \\ \text{L} \end{array} \}$$

$\leftarrow \quad \uparrow \quad \downarrow$   
 $0.8 \times (1) + 0.1 \times (0.7) + 0.1 \times (-1) \quad (L) \quad (0.76)$

U & D CAN'T WIN

(b) For each state in the figure above, draw an arrow indicating the optimal policy following the 2nd iteration of value iteration.

See figure, ties broken alphabetically in  $S_1$

(c) Now assume that the robot is moving on an infinite plane with no obstacles or goals, and the reward at each state is 0.5. Using a discounting factor of 0.9, what is the expected utility of the optimal policy?

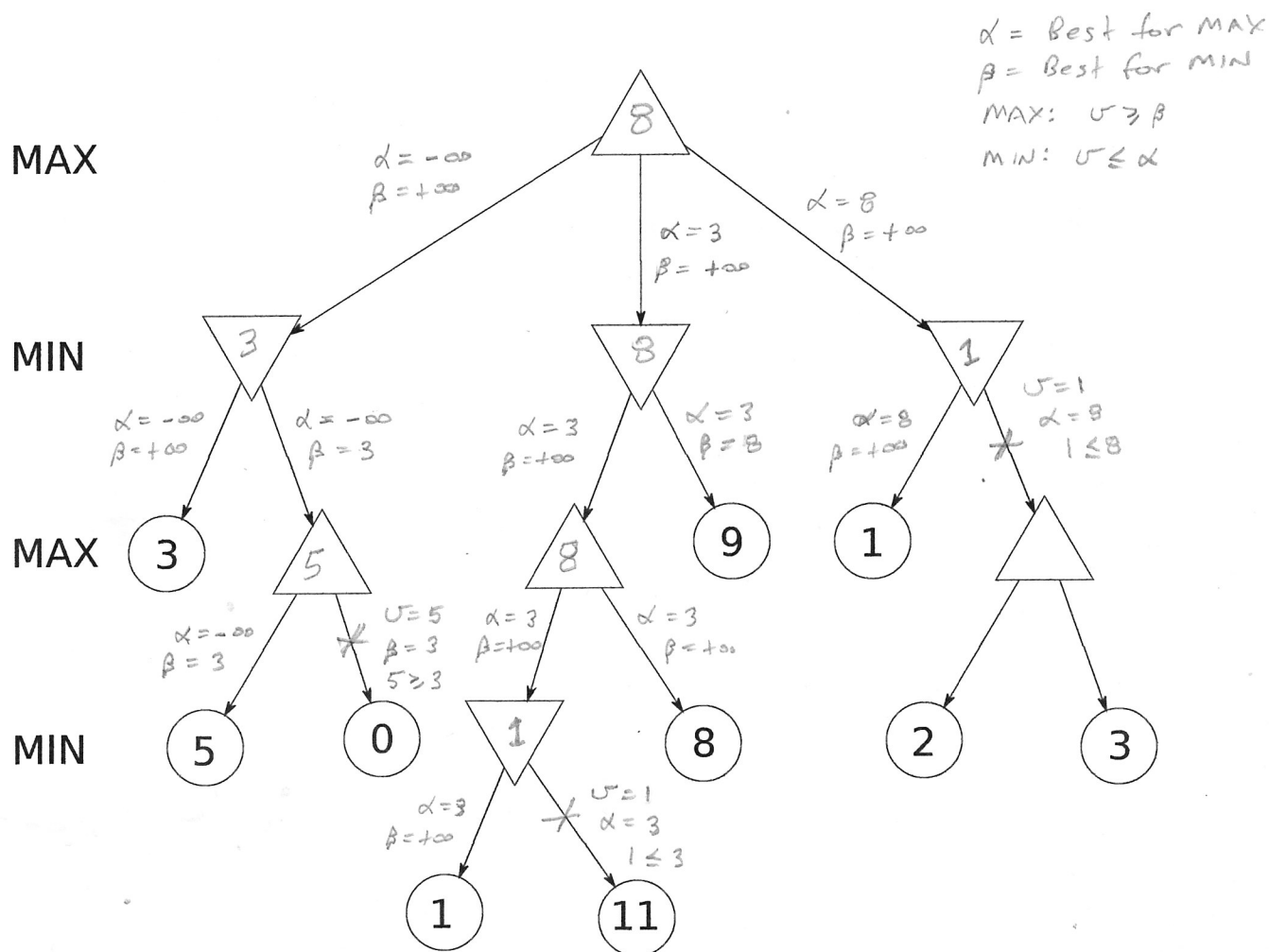
The question is asking about infinite horizon case

$$U = \sum_{t=0}^{\infty} \gamma^t R(S_t) = 0.5 \sum_{t=0}^{\infty} (0.9)^t = \frac{0.5}{1-0.9} = 5$$

Note: Since all directions equal, all actions equal. Then  $t$  is just how many cells visited

See (17.1) in book

**Question 7.** Consider the following tree for a minimax game. Note: End utilities given are always with respect to MAX, even if they don't appear on a MAX layer.



Note: The  $\alpha, \beta$  values on each branch are the values that are passed in to the function call for the node.  $\alpha$  can only change across the actions for a MAX node.  $\beta$  can only change across the actions for a MIN node.

Use the **Minimax algorithm with Alpha-Beta pruning** to identify the best strategy for MAX. Assume the algorithm always searches successors from left to right. Cross out all branches that are pruned. For each prune, write the value of  $v$  and  $\alpha$  or  $\beta$  (whichever applies to the pruning decision). Fill in the value of every node, as the algorithm would determine.

**Question 10.** (Adapted from problem 8.28 in the book)

Consider a first order logic knowledge base that describes worlds containing people, songs, and albums. The vocabulary contains the following symbols:

$Sings(p,s,a)$ : Album  $a$  includes a recording of a song  $s$  sung by a person  $p$

$Wrote(p,s)$ : Person  $p$  wrote song  $s$

Express the following sentences in first order logic

- Either Gershwin or McCartney wrote "The Man I Love"

$$(wrote(G, "TMIL") \vee wrote(M, "TMIL")) \wedge \neg (wrote(G, "TMIL") \wedge wrote(M, "TMIL"))$$

- Gershwin never recorded (i.e. sang) "The Man I Love"

$$\forall a \neg sings(G, "TMIL", a)$$

- Gershwin did not write any of the songs on "Revolver"

$$\forall s, p \ sings(p, s, "Revolver") \Rightarrow \neg wrote(G, s)$$

- There is a single album containing every song that McCartney has written

$$\exists a (\forall s \ wrote(M, s) \Rightarrow \exists p \ sings(p, s, a))$$