/

COMP4200/5430: Artificial Intelligence

Exam II – 23rd April 2019

Name: Hoang Do

Notes are not allowed. Calculators are allowed.
Note: If something is not clear, state your assumption, answer the question and move on.

You have 75 minutes

Problem	**Topic**	**Possible Points**	**Your Score**
1	Bayes Theorem	10	3
2	Bayes Nets	15	13
3	MDP_I	15	15
4	MDP_II	10	8
		50	39

1 Bayes Theorem (10 Points)

Assume that Alex wears a kilt about once a year, and Jared wears a kilt about once every five years. Also, assume that Jared walks down a particular sidewalk every day, but Alex only walks down the sidewalk every other day[1].

i. (2 points) If you see a person walking down the sidewalk in the distance, and you are sure that he is either Alex or Jared, but you can't tell which, and you also can't tell what he is wearing, what are the following probabilities?

P(Alex): 0.5 1/3
P(Jared): 0.5 2/3

ii. (8 points) Now suppose the person is wearing a kilt. What is the probability that it is Alex? (show your calculations)

$$P(Alex \mid Kilt) = \frac{P(Kilt \mid Alex) . P(Alex)}{P(Kilt)}$$

$$= \frac{P(Kilt \mid Alex) . P(Alex)}{P(Kilt \mid Alex) . P(Alex) + P(Kilt \mid Jared) . P(Jared)}$$

3

$$= \frac{1 . 0.5}{1 . 0.5 + \frac{1}{5} . 0.5} = 0.83$$

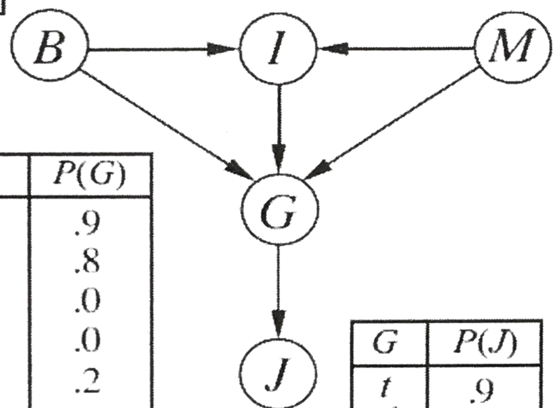
[1] Every other day means once every two days.

2 Bayesian Networks & D-Separation (15 points)

B	M	P(I)
t	t	.9
t	f	.5
f	t	.5
f	f	.1

P(B)
.9

P(M)
.1



B	I	M	P(G)
t	t	t	.9
t	t	f	.8
t	f	t	.0
t	f	f	.0
f	t	t	.2
f	t	f	.1
f	f	t	.0
f	f	f	.0

G	P(J)
t	.9
f	.0

I. (5 points) Use d-separation to determine the following (write True or False):

F a) $B \perp M \mid G$

F b) $B \perp M \mid I$

~~F~~ c) $B \perp J \mid G$ True

F d) $B \perp G$

~~T~~ e) $B \perp M \mid J$ False

3

II. (5 points) Calculate the value of P(b, i, ¬m, g, j).

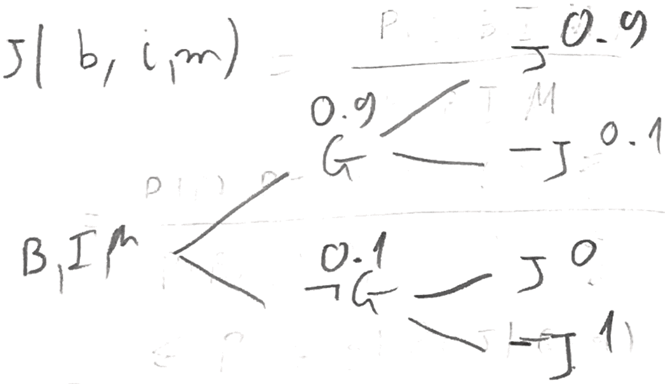
$P(b, i, \neg m, g, j) = P(b) . P(\neg m) . P(i | b, \neg m) . P(g | b, i, \neg m) . P(j | g)$

$= 0.9 \times 0.9 \times 0.5 \times 0.8 \times 0.9$

$= 0.2916$

III. (5 points) Calculate the probability that someone goes to jail given that they broke the law, have been indicted, and face a politically motivated prosecutor.

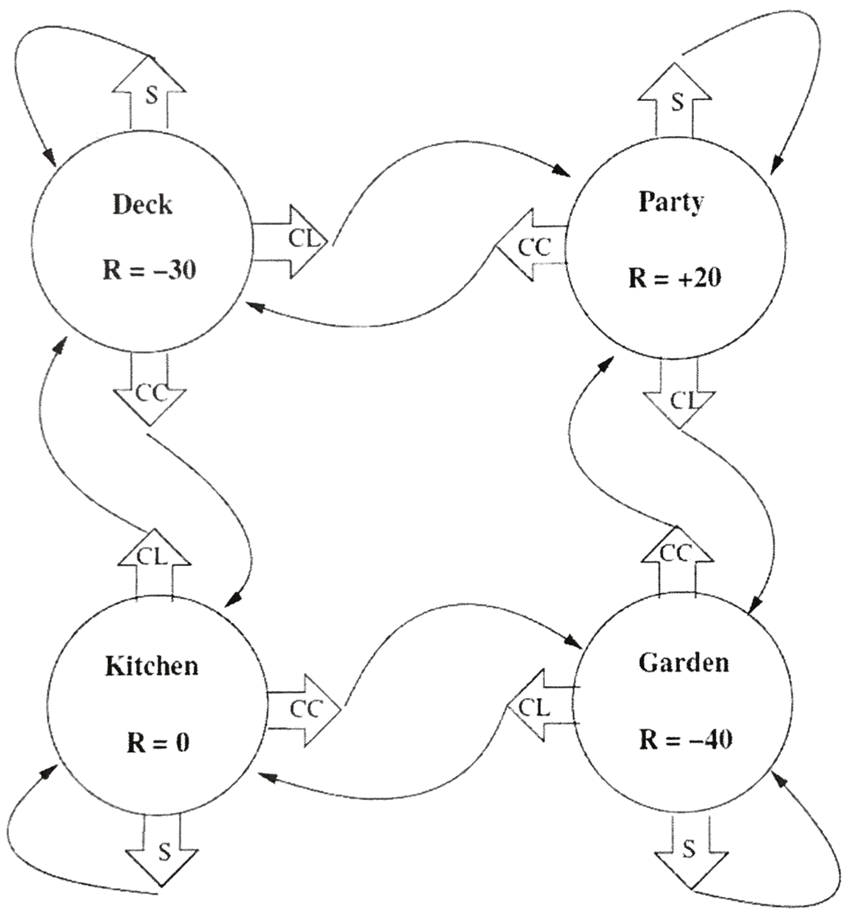
$P(J | b, i, m) =$



$\Rightarrow P(J | b, i, m) = 0.9 \times 0.9 + 0.1 \times 0 = 0.81$

3 MDPS_I (15 Points)

You are a wildly implausible robot who wanders among the four areas depicted below. You hate rain and get a reward of -30 on any move that starts in the deck (D) and -40 on any move that starts in the Garden (G). You like parties (P), and you are indifferent to kitchens (K)



Actions: All states have three actions: Clockwise (**CL**), Counter-Clockwise (**CC**), Stay (**S**). Clockwise and Counter-Clockwise move you through a door into another room, and Stay keeps you in the same location. All transitions have are deterministic (probability 1.0).

a) (1 point) How many distinct policies are there for this MDP?

$3^4 = 81$.

1

b) (8 points) Let V*(Room)= expected discounted sum of future rewards assuming you start in "Room" and subsequently act optimally. Assuming a discount factor γ=0.5, give the V* values for each room.

$$\left.\begin{aligned} V^*(D) &= -30 + 0.5 \times V^*(P) \\ V^*(P) &= +20 + 0.5 \times V^*(P) \\ V^*(K) &= 0 + 0.5 \times V^*(K) \\ V^*(G) &= -40 + 0.5 \times V^*(P) \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} V^*(D) &= -10 \\ V^*(P) &= 40 \\ V^*(K) &= 0 \\ V^*(G) &= -20 \end{aligned}\right.$$

8

c) (6 points) The optimal policy when the discount factor, γ, is small but non-zero (e.g. γ=0.1) different from the optimal policy when is large e.g. γ=0.9). If we began with γ=0.1, and then gradually increased, what would be the threshold value of above which the optimal policy would change? (Show calculations and justify your answers).

If the discount factor γ increase. the policy of Kitchen change from S to CL. This happen at the value of taking action S equal to CL 2

$$V^S(\text{Kitchen}) = 0 + \gamma V^{*S}(\text{Kitchen}) = 0 \quad 1$$

$$V^{CL}(\text{Kitchen}) = 0 + \gamma V^*(\text{Deck}) \quad 1$$

We have

$$V^*(\text{Deck}) = -30 + \gamma V^*(\text{Party})$$

$$V^*(\text{Party}) = 20 + \gamma V^*(\text{Party})$$

$$\Rightarrow -30 + \gamma \left(\frac{20}{1-\gamma}\right) = 0 \quad 2$$

$$\Rightarrow \gamma = 0.6$$

4 MDPS_II (10 Points)

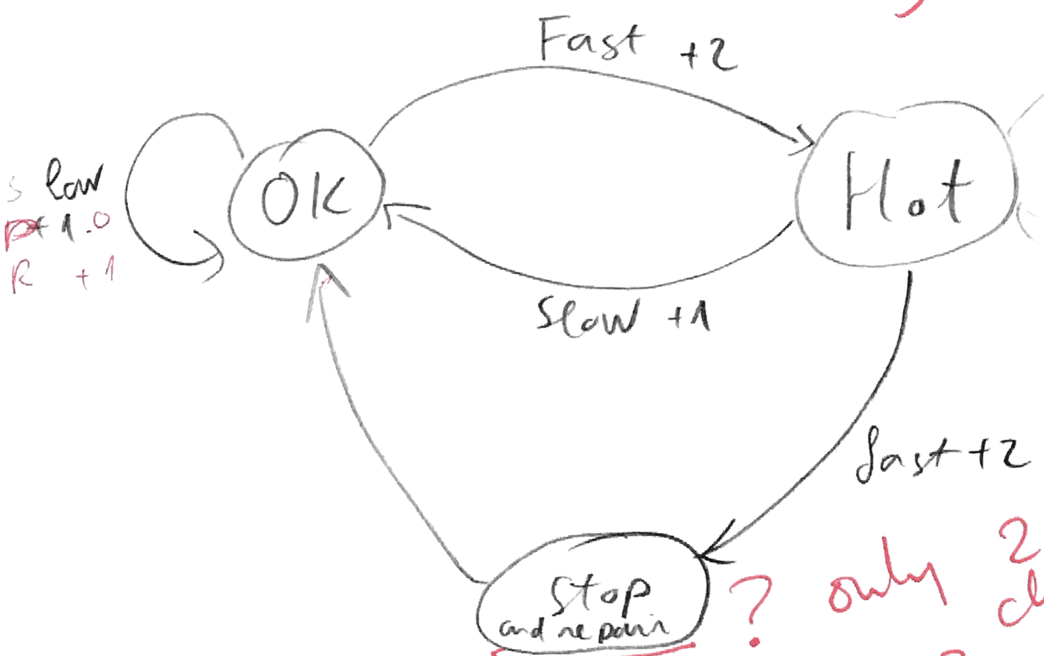
Consider an autonomous robot which can either move or **FAST** or **SLOW** in any time step. Moving fast generally gives a reward of +2 while moving slowly gives a reward of only +1. However, the robot must also take into account its internal temperature which can either be **HOT** or **OK**. Driving **SLOW** lowers the temperature while driving **FAST** tends to raise it. If the robot is HOT then there is a danger of overheating at which time it must then stop, cool down and make repairs. The MDP Transitions and Rewards are specified as follows:

s	**a**	**s'**	**T(s,a,s')**	**R(s,a,s')**
OK	SLOW	OK	1.0	+1
OK	FAST	OK	0.5	+2
OK	FAST	HOT	0.5	+2
HOT	SLOW	OK	1.0	+1
HOT	FAST	HOT	0.5	+2
HOT	FAST	OK	0.5	-10

Note: while repairs are costly, the robot is OK afterwards (last row)

a. **(5 points)** Draw a decision network and correctly label the states, action, transition and reward models etc.

3



? only 2 states. check above.
-2.

b. **(5 points)** Run two rounds of value iteration in the table below using $\gamma = \mathbf{0.8}$. You may skip the greyed out cell.

S	V_0	V_1	V_2
OK	0	2	3.2
HOT	0	1	

$V_1(OK) = \max[1(1+\gamma \cdot 0), 0.5(2+\gamma \cdot 0) + 0.5(2+\gamma \cdot 0))] = 2$

good

$V_1(hot) = \max[1(1+\gamma \cdot 0), 0.5(2+\gamma \cdot 0) + 0.5(-10+\gamma \cdot 0)] = 1.$

$V_2(ok) = \max[1(1+\gamma \times 2), 0.5(2+\gamma \times 2) + 0.5(2+\gamma \cdot 1)]$

$= 3.2.$