Artificial Intelligence : DS251/CSL304

EndSem

Time - 3 hrs

Full Marks 65

1. The monkey-and-bananas problem is faced by a monkey in a laboratory with some bananas hanging out of reach from the ceiling. A box is available that will enable the monkey to reach the bananas if he climbs on it. Initially, the monkey is at A, the bananas at B, and the box at C. The monkey and box have height Low, but if the monkey climbs onto the box he will have height High, the same as the bananas. The actions available to the monkey include Go from one place to another, Push an object from one place to another, ClimbUp onto or ClimbDown from an object, and Grasp or Ungrasp an object. The result of a Grasp is that the monkey holds the object if the monkey and object are in the same place at the same height. The goal here is to make the monkey grasp the bananas hanging from the ceiling.

a. Write down the initial state description.

b. Write the six action schemas.

c. Apply the partial-order planning algorithm to draw the final partial order plan clearly showing the causal links (including the subgoals that they achieve) and ordering links.

[2+3+7=12 Marks]

2. a. Consider the following Car Tire replacement planning problem. Draw the GraphPlan for the first layer and show all the mutex links.

Start: At(Flat, Axle) ∧ At(Spare, Trunk)
Goal: At(Spare, Axle)

Op(ACTION: Remove(Spare, Trunk), PRECOND: At(Spare, Trunk), EFFECT: At(Spare, Ground) ∧ ¬ At(Spare, Trunk))

Op(ACTION: Remove(Flat, Axle), PRECOND: At(Flat, Axle), EFFECT: At(Flat, Ground) ∧ ¬ At(Flat, Axle))

Op(ACTION: PutOn(Spare, Axle), PRECOND: At(Spare, Ground) ∧ ¬ At(Flat, Axle), EFFECT: At(Spare, Axle) ∧ ¬ At(Spare, Ground))

Op(ACTION: LeaveOvernight, PRECOND: None, EFFECT: ¬ At(Spare, Ground) ∧ ¬ At(Spare, Axle) ∧ ¬ At(Spare, Trunk) ∧ ¬ At(Flat, Ground) ∧ ¬ At(Flat, Axle))

b. In GraphPlan, explain whether the following statements are true or false with justifications - (i) Literals increase monotonically; (ii) Actions increase monotonically; (iii) Mutexes increase monotonically.

[7+3=10 Marks]

3. Consider the following axioms,

If the alarm sounds, then there is a fire or there is a drill.
If there is a drill, then the fire department will not come.
The fire department comes.
Therefore, if the alarm sounds, then there is a fire.

Represent all of them in Propositional Logic. Prove the conclusion using the resolution refutation method. Show all required steps.

[6 Marks]

4. a. Consider two medical tests, A and B, for a virus. Test A is 95% effective at recognizing the virus when it is present, but has a 10% false positive rate (indicating that the virus is present, when it is not). Test B is 90% effective at recognizing the virus, but has a 5% false positive rate. The two tests use independent methods of identifying the virus. The virus is carried by 1% of all people. Say that a person is tested for the virus using only one of the tests, and that test comes back positive for carrying the virus. Which test returning positive is more indicative of someone really carrying the virus? Justify your answer mathematically.

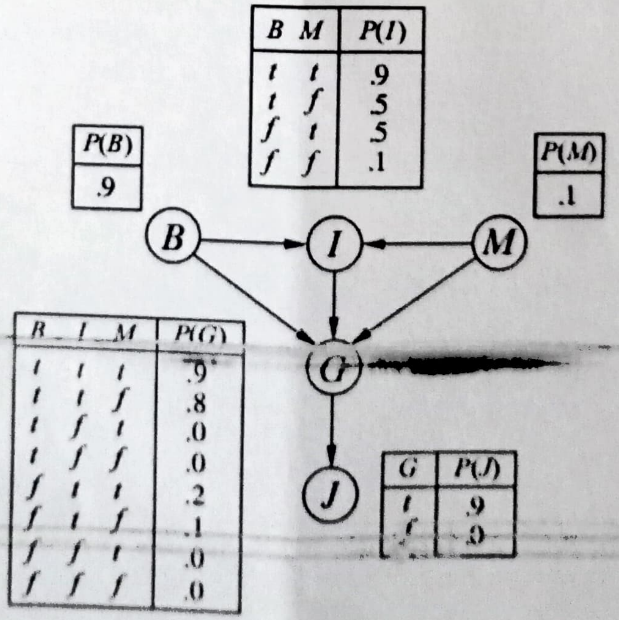
b. Let's consider two fuzzy sets A and B defined on the same universe X={1,2,3,4,5} where the corresponding membership functions defined as

μ_A={(1,0.2),(2,0.5),(3,0.7),(4,0.4),(5,0.1)} and μ_B={(1,0.6),(2,0.3),(3,0.4),(4,0.8),(5,0.2)}.

Find $\mu_{A\cap B}(x)$ and $\mu_{A\cup B}(x)$.

[5+3 = 8 Marks]

5. Consider the following bayesian belief network,



P(B)
.9

P(M)
.1

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

Where boolean variables B indicates "BrokeElectionLaw" , I indicates "Indicted (charged for a crime)", M indicates "PoliticallyMotivatedProsecutor" , G indicates "FoundGuilty", and J indicates "Jailed" .

a. Which of the following are asserted by the network structure? Justify your answer.

(i) P(B, I, M) = P(B)P(I)P(M).

(ii) P(J | G) = P(J | G, I).

(iii) P(M | G, B, I) = P(M | G, B, I, J).

b. Calculate the value of P (B, I, ¬M, G, J).

c. Calculate the probability that someone goes to jail given that they broke the law, have been indicted, and faced a politically motivated prosecutor.

d. A context-specific independence allows a variable to be independent of some of its parents given certain values of others. In addition to the usual conditional independences given by the graph structure, what context-specific independences exist in the Bayes net in the Figure above?

[3+2+3+2=10 Marks]

6. Consider an undiscounted MDP having three states, (1, 2, 3), with rewards −1, −2, 0, (can also be considered as initial values of the states) respectively. State 3 is a terminal state. In states 1 and 2 there are two possible actions: a and b. The transition model is as follows:

a. In state 1, action a moves the agent to state 2 with probability 0.8 and makes the agent stay put with probability 0.2.

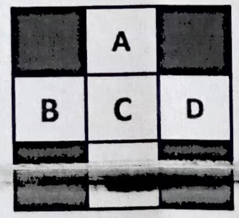
b. In state 2, action a moves the agent to state 1 with probability 0.8 and makes the agent stay put with probability 0.2.

c. In either state 1 or state 2, action b moves the agent to state 3 with probability 0.1 and makes the agent stay put with probability 0.9.

Apply policy iteration, showing each step in full, to determine the optimal policy and the values of states 1 and 2. Assume that the initial policy has action b in both states.

[8 Marks]

7. Consider the following Gridworld example. We would like to use TD learning to find the values of these states.



Suppose we observe the following (s, a, s', R(s, a, s')) transitions and rewards:
(B, East, C, 2), (C, South, E, 4), (E, North, A, 4), (C, East, A, 6), (B, East, C, 2)

Note that the R(s, a, s') in this notation refers to observed reward, not a reward value computed from a reward function. The initial value of each state is 0. Let, discount factor $\gamma = 1$ and learning rate $\alpha = 0.5$.

What are the learned values for each state from TD learning after all five observations/samples?

[4 Marks]

8. Consider a Markov Decision Process (MDP) with two states $S=\{s_1,s_2\}$ and two actions $A=\{a_1,a_2\}$. The discount factor is $\gamma=0.8$. You are given the following information for a single transition:

- Current state: $s_t=s_1$
- Action taken: $a_t=a_1$
- Reward received: $r_t=5$
- Next state: $s_{t+1}=s_2$
- Current Q-values: $Q(s_1,a_1)=2$, $Q(s_2,a_1)=4$, $Q(s_2,a_2)=1$
- Learning rate: $\alpha=0.5$

Compute the updated value of $Q(s_1,a_1)$ after the transition.

b. Give an example situation where we cannot use traditional tabular Q-Learning and have to use Deep Q Learning.

c. Explain how exploration functions stop exploring actions whose badness is established.

[3+2+2=7 Marks]