

## Primary Examination, Semester 1, 2016

<b>Artificial Intelligence COMPSCI 3007, 7059</b>
---

Official Reading Time:	10 mins
Writing Time:	120 mins
Total Duration:	130 mins

Questions	Time	Marks
Answer all 5 questions	120 mins	120 marks
		120 Total

### Instructions

- Begin each answer on a new page in the answer book.
- Examination material must not be removed from the examination room.

### Materials

- No calculators allowed. Answers may be left in the form of numerical expressions.
- Foreign language paper dictionaries permitted.

DO NOT COMMENCE WRITING UNTIL INSTRUCTED TO DO SO

## Problem solving by searching

## Question 1

The following is a map of a country with several cities connected by highways. The numbers attached to the lines connecting any two cities denote the length of the connecting highways.

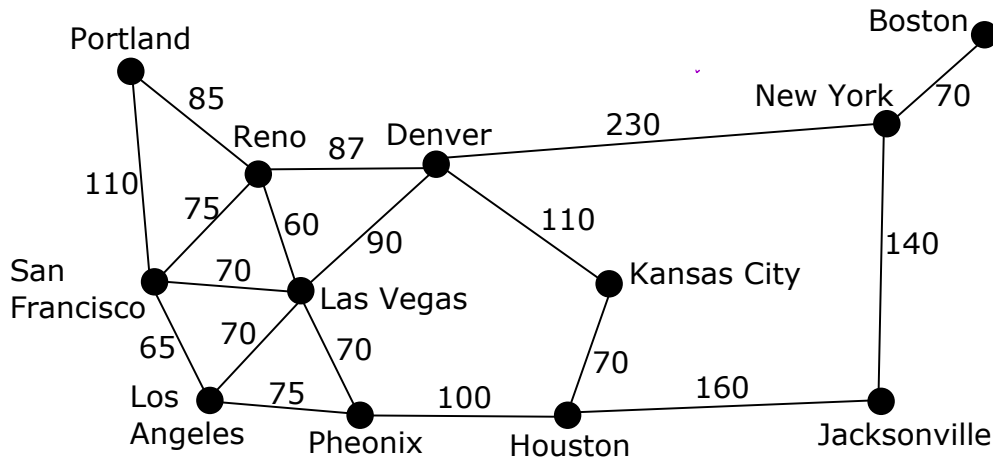


Figure 1: Map of a country.

- (a) We aim to find a path from Portland to Boston. Draw the search tree that solves this problem based on Uniform Cost Search (UCS). Indicate clearly the path cost at each node. Remember to avoid repeated states.

[6 marks]

- (b) In the context of pathfinding, define the start distance of a given state.

*from start point to its direct neighbours' distance.*

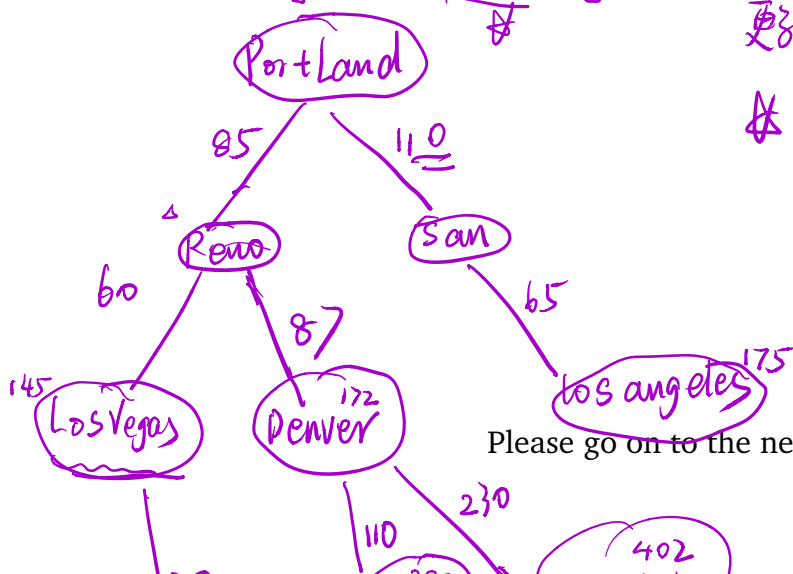
[2 marks]

- (c) Copy the map in Figure 1 to your answer booklet. Then, using Portland as the start state, calculate the start distances of all the cities, and annotate the start distances clearly on the map.

[6 marks]

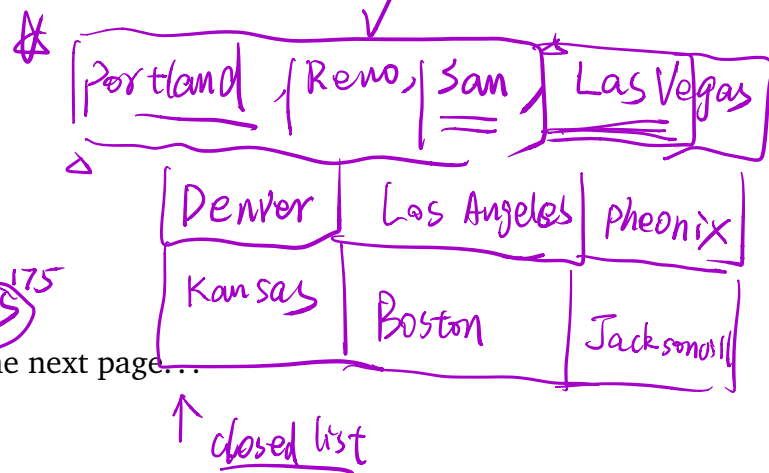
- (d) Given the start distances of all the states, how would you find the desired shortest path?

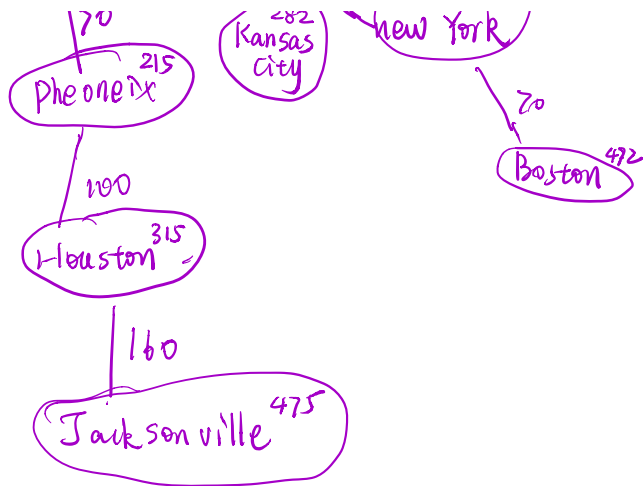
[2 marks]



Please go on to the next page...

Don't forget





Start out empty ->

- (e) Changes have occurred to the map in Figure 1. The new map is shown in Figure 2, where changes are highlighted in red. With reference to the start distances you have calculated in part (c) above, which states are now locally inconsistent? Show your workings clearly.

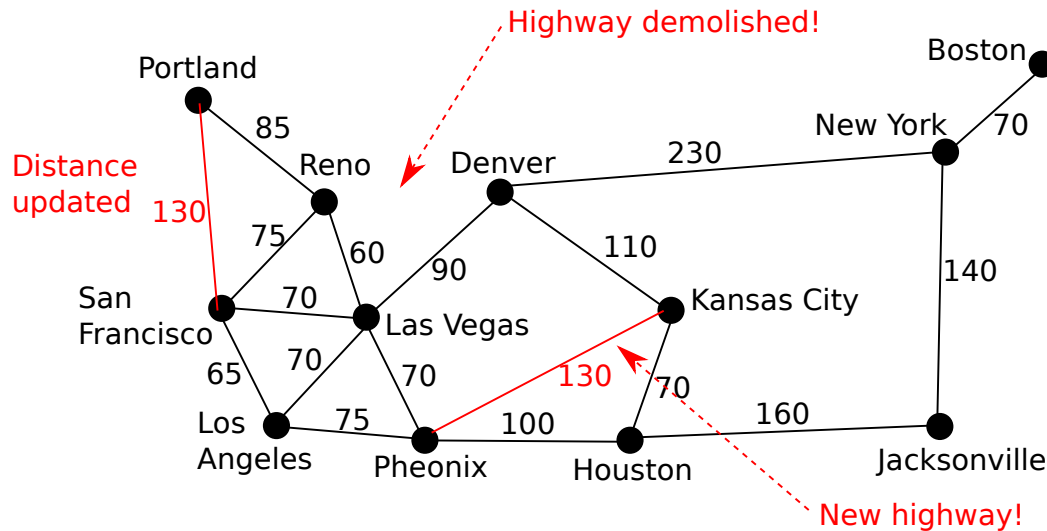


Figure 2: New map with changes highlighted in red.

[6 marks]

[Total for Question 1: 22 marks]

## Game playing I

## Question 2

- (a) What does it mean by perfect information in the context of two-person adversarial games?

each player knows the initial and all the history event of the game.

[1 mark]

- (b) When applying the minimax algorithm on the game of checkers, what game rule must be imposed to ensure that the game tree is finite?

finite positions and checkers can ensure the game tree is finite

[2 marks]

- (c) The following figure shows a game tree of up to three moves. The first move belongs to Max.

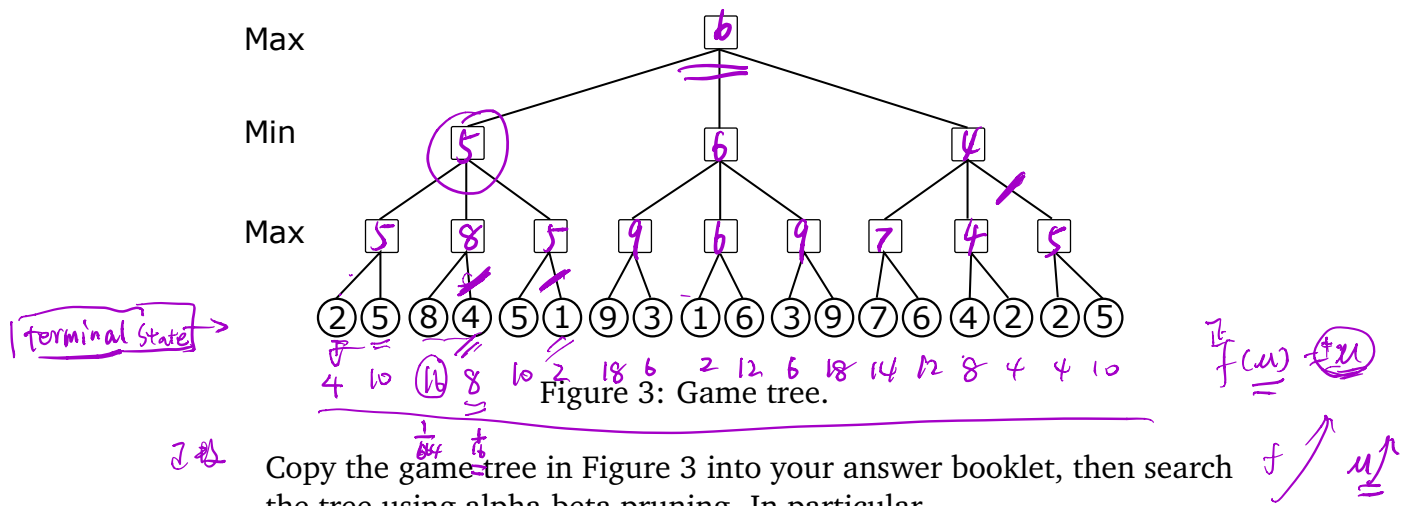


Figure 3: Game tree.

Copy the game tree in Figure 3 into your answer booklet, then search the tree using alpha-beta pruning. In particular,

- Write clearly the minimax value at the root node; and
- Circle clearly the branches that are pruned (i.e., branches that do not need to be visited).

[3 marks]

- (d) The utility values in the game tree in Figure 3 are to be updated according to a function  $f(u)$ , where  $u$  is the old utility value.

- i. Under which of the following 5 update functions  $f$ , will lead to the same branches being pruned by alpha-beta pruning?

1.  $f(u) = 2u^2 + 3$  ✓

2.  $f(u) = (u - 3)(u - 7)$  ✗

3.  $f(u) = 1/u^2$  ✗

4.  $f(u) = 3 \ln(u) + \tan(5\pi/4)$  ✗

5.  $f(u) = \min(u^3, \cos(u))$  ✗

- ii. Justify your answer above.

✓  $x \geq 1 \rightarrow f(u) = 4u$ , 单调增,  $\rightarrow$  相符

$f'(u) = 2u - 10$ , 非单调增,

$f'(u) = -2u^{-3}$ ,  $\rightarrow$  单调减  $\rightarrow$  不符

$f'(u) = 3/u$ , 在  $u > 0$  区间增,  $\rightarrow$  ✓

$u > 1, u^3 \neq \cos(u) \sim (-1, 1)$

$\min(u^3, \cos(u)) = \cos(u)$

$\cos(u)$  不单调 不符

[6 marks]



(e) If you are allowed to reorder the <sup>叶节点</sup>leaf nodes in the original game tree in Figure 3 before alpha-beta pruning is conducted,

- i. How would you reorder the leaf nodes so as to increase the number of branches that can be pruned?
- ii. Draw the game tree with your suggested ordering in part i above, and clearly circle the branches that are pruned.

[6 marks]

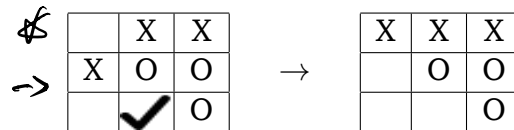
[Total for Question 2: 18 marks]

✓  
Game playing II

## Question 3

Three Men's Morris is a played on a 3x3 board. Each of the two players is allowed three game pieces. The objective is to place the three pieces on a straight line (horizontally or vertically) on the board. The game begins with a blank board, and the players take turns to place the pieces one-by-one on the board. If the game has not ended by then, the players take turns to move the pieces until one of them wins. Each piece can be moved horizontally or vertically (but not diagonally) to an adjacent empty location.

The left diagram below shows an example board state, where the move belongs to Max. If Max moves the leftmost X piece up, it will lead to the board state in the right diagram, which corresponds to an end state.

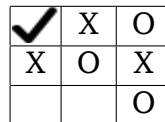


- (a) What is the branching factor of Three Men's Morris?

3

[5 marks]

- (b) Starting from the board state below, where the move belongs to Max, draw the game tree up to two plies (one move for each player).



[8 marks]

- (c) For each of the leaf nodes in the partial game tree that you have drawn in part (b) above, evaluate its utility value according to the function

$$\text{Longest X} - \text{Longest O}$$

where "Longest X" is the longest vertical or horizontal sequence of X in the board state, and similarly for "Longest O". Annotate the calculated utility values in the game tree clearly.

[6 marks]

- (d) Based on the game tree that you have drawn in part (b) and the evaluation value calculated in part (c), what is the best move to make at the board state in part (b) according to the minimax principle?

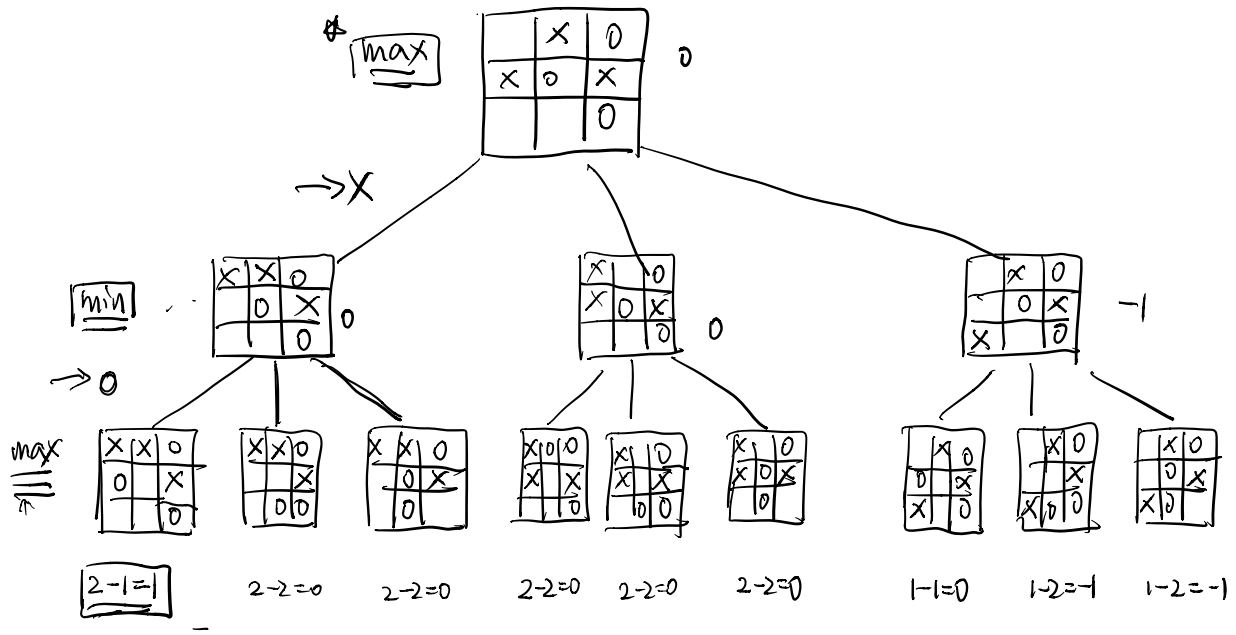
[2 marks]



[Total for Question 3: 21 marks]

Please go on to the next page...





## Learning

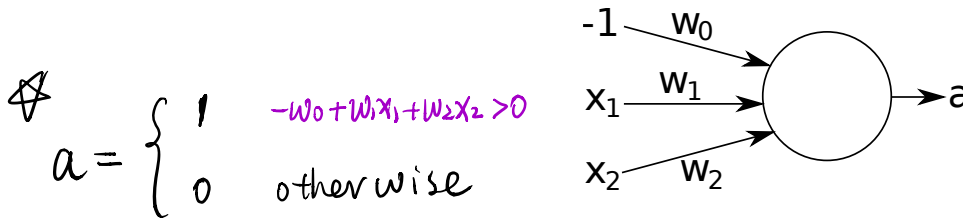
## Question 4

- (a) In the context of learning systems, what is the principle of Occam's Razor?

奥卡姆的剃刀, 若无必要, 不增加复杂度.

[1 mark]

- (b) The following shows a perceptron unit with two inputs.



- Write an expression for the output as a function of the other quantities in the unit.
- Explain why the perceptron can only solve linearly separable problems.

[2 marks]

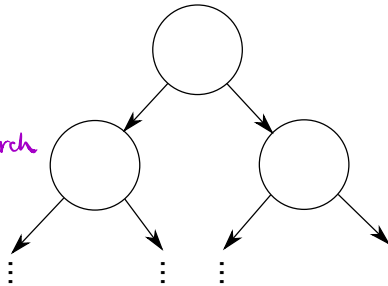
- (c) Figure 4 in Page 8 shows a set of points on the plane.

- i. Construct a kd-tree for this set of points, using the split rule  $x \leq \text{splitval}$  or  $y \leq \text{splitval}$ . Draw your kd-tree in the following format, and remember to indicate clearly the split rule at each node.

k-dimensional tree

k维里组数据树.

range and nearest neighbour search



Hint: Each leaf node in your kd-tree should correspond to a point in Figure 4.

[6 marks]

- ii. We wish to find the nearest neighbour of query point (5,81) using the kd-tree search algorithm. Which are the points in Figure 4 that need to be compared with the query point? List the coordinates of those points here.

[6 marks]



- (d) The following is a set of observations of 10 prominent computer scientists: whether they have a Maths degree, whether they have won the Turing Award, whether they drive a BMW, and the number of papers they publish per year (20, 30 or 40).

Scientist	Maths	Turing	BMW	Papers
s1	N	N	Y	20
s2	N	N	N	20
s3	N	N	N	20
s4	Y	N	N	30
s5	Y	N	N	30
s6	Y	Y	Y	40
s7	N	Y	Y	40
s8	N	Y	N	40
s9	Y	Y	Y	30
s10	Y	Y	Y	30

We aim to build a decision tree that can predict whether an early career computer scientist will win a Turing Award in his/her lifetime, based on the other attributes. To answer the following questions, these numbers will be useful:  $\log_2 0.1 = -3.32$ ,  $\log_2 0.2 = -2.32$ ,  $\log_2 0.3 = -1.73$ ,  $\log_2 0.4 = -1.32$ ,  $\log_2 0.6 = -0.73$ ,  $\log_2 0.8 = -0.32$ .

- i. What is the information content at the root node?
- ii. Calculate the information gain of attributes Math and BMW at the root node.
- iii. Between Math and BMW, which is the better attribute to use at the root node?

[2 marks]

[6 marks]

[2 marks]

**[Total for Question 4: 25 marks]**

**Probabilistic reasoning**

This is called a first order Markov assumption since we say that the probability of an observation at time  $n$  only depends on the observation at time  $n-1$

**Question 5**

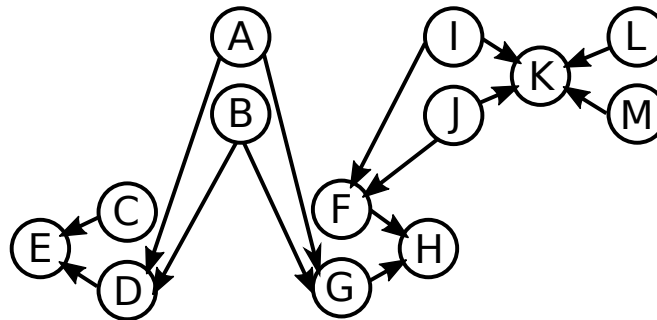
(a) Why is it necessary to consider approximate inference methods?

[1 mark]

(b) In the context of probabilistic reasoning over time, what does it mean by the first order Markov assumption?

[1 mark]

(c) The following is a Bayesian network of 13 Boolean variables.



What is the number of independent probability values required to fully define the Bayesian network?

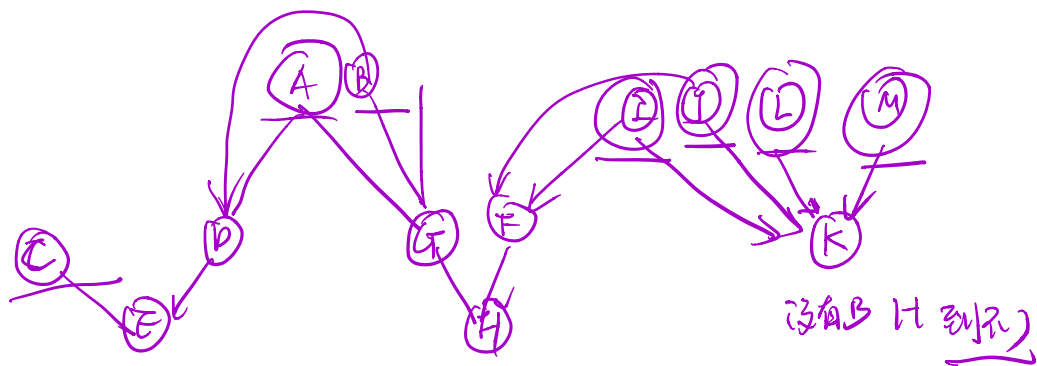
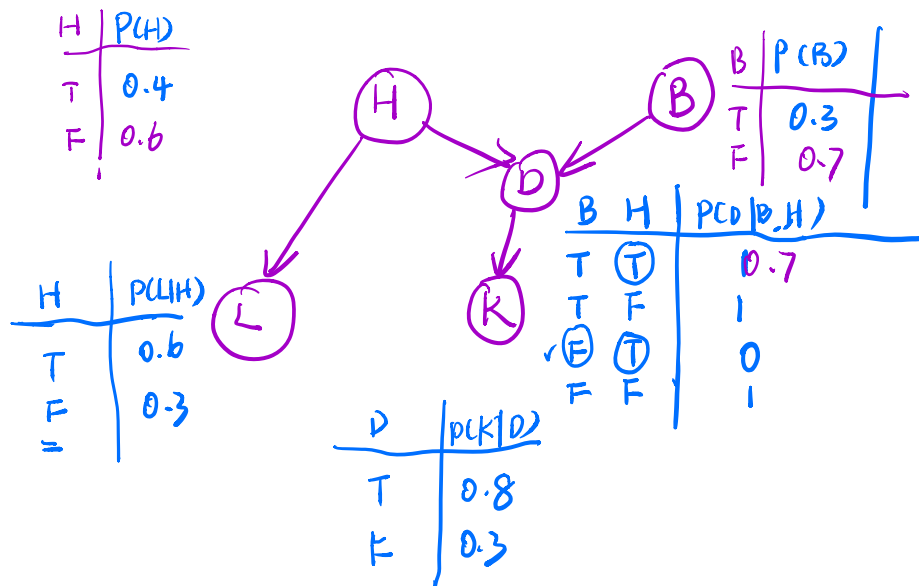
[4 marks]

(d) This following description is adapted from Charniak, 1991:

I share a house with my brother. When I go home, I want to know if anyone is home before I go in. From past experience, when my brother leaves the house, with probability 0.3 he turns on the outside light. When he is at home, sometimes he still turns the outside light on since he is expecting guests. When nobody is home, our dog is always left outside. When my brother is home, the dog is never left outside, unless if it is suffering from bowel-troubles, in which case it is left outside with probability 0.7. If the dog is outside, I will hear it barking 80% of the time. Sometimes, when our dog is not outside, with probability 0.3 I still hear barking. My brother is not at home 60% of the time when I go home, and the dog has bowel-troubles with probability 0.3 at any day. My brother expects guests 60% of the time when he is at home.

The variables in the above system are represented by the following symbols:

- H: my brother is at home.
- L: the outside light is on.
- D: the dog is outside.
- B: the dog has bowel-troubles. ✓
- K: I can hear barking.



$\mathbb{Z}^+$  没有

Given H

- H: my brother is at home. ✓
- L: the outside light is on. ✓
- D: the dog is outside. ✓
- B: the dog has bowel-troubles. ✓
- K: I can hear barking. ✓

i. Draw the Bayesian network corresponding to the above description. Remember to define the relevant CPTs. ✓

[12 marks]

✓ ii. What is the global semantic (i.e., joint distribution of all the variables) encoded by the Bayesian network?

$$P(H, L, D, B, K) \rightarrow P(H) \cdot P(L|H) \cdot P(K) \cdot P(D|B, H) \cdot P(B)$$

[2 marks]

$$P(\neg H | K, L)$$

iii. If I can hear barking and can see that the outside light is on when I get home, what is the probability that nobody is at home?

[4 marks]

(e) A computer science student eats one of three types of food for dinner: salad, pasta, or burger - in decreasing level of healthiness. What the student will eat today depends on what he had yesterday. Specifically, there is a 10% chance that the student will have the same thing as yesterday. If the student decides to have something different than yesterday, he will choose the healthier option twice as likely as the less healthy one. At any given night, the trash can of the student will consist of the container for his dinner: either a plastic box or a tin foil wrapper. It is known that, in 7 out of 10 times, salad and pasta are sold in plastic boxes. 80% of the time, burgers are wrapped in tin foil. 2

i. The eating habit above can be summarised by a HMM. Define the state transition matrix T if this HMM.

[4 marks]

ii. The mother of the student wishes to monitor what he has been eating, without asking him directly. She made a record of the type of container in his trash can for several days.

day	1	2	3	4	5
container	box	foil	foil	box	box

Given the observations above, derive the probability distribution of the type of food F consumed by the student on

- day 1; and
- day 2.

Leave the probability distributions as numerical expressions, and without solving for any normalising constants. Assume the student was equally likely to consume the food types on day 0.

[6 marks]

[Total for Question 5: 34 marks]

	→	salad	pasta	burger
→ salad		0.10	0.6	0.3
✓ pasta		0.6	0.1	0.3
✓ burger		0.6	0.3	0.1

observation (box)

salad	0.7		
paste		0.7	
burge			0.2

foil

0.3		
	0.3	
		0.8