BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
CSF 407 (ARTIFICIAL INTELLIGENCE), Second Semester 2021 - 2022
Comprehensive Exam [CLOSED BOOK]

May 10, 2022 **MM: 38**

Answer all parts of a question together. Elaborate your answers wherever required.

Q1 [4M]. Answer the following:

a) Consider that you have to design the application of "*Speech to Text Conversion*". Assuming that the input can be noisy, suggest a method to correct the noise using probability distribution (or probability knowledge). Support your answer using mathematical formulations and assumptions wherever possible. Limit your answer to the topics studied during the course.

b) The following list of R's and N's represents relevant (R) and non-relevant (N) returned documents in a ranked list of 20 documents retrieved in response to a query from a collection of 10,000 documents. The top of the ranked list (the document the system thinks is most likely to be relevant) is given below. This list shows 6 relevant documents, and rest non-relevant. Assume that there are 8 relevant documents in total in the collection. What is the precision and recall of the system on the top 20?

R R N N N N N N R N R N N N R N N N N R

Q2 [4M]. We can improve the predictions made by a rational agent by learning from data. The task is to learn a hypothesis that fits the future data best. Answer the following:

a) Is there any restriction on future data?

b) Suppose we want to fit a polynomial of appropriate degree on the given data. *Minimizing squared error loss is always a good metric to select the hypothesis*. True/False. If True, explain squared error loss and why it is superior over absolute value loss. If False, suggest a method of selecting the hypothesis.

c) Suppose L is the loss function (absolute/squared/any-other) and E = (X, Y) be the set of input-output examples. Then the expected generalization loss for a hypothesis h is:

$$\text{GLOSS (h)} = \sum_{(x,y)\in E} L\big(y, h(x)\big)P(x,y).$$

Can we always accurately find the minimum expected generalization loss? Justify your answer.

Q3 [4M]. Nearest Neighbour models works very well in all situations. True/False. Justify and prove your answer.

Q4 [2M] With respect to the set of training examples shown in Table-1, find

a) the entropy of the given set of training examples, and

b) the information gain of X2 relative to these training examples.

Instance	Classification	X1	X2
1	+	T	T
2	+	T	T
3	-	T	F
4	-	F	F
5	-	F	T
6	-	F	F

Table-1

Q5 [4M]. A patient goes to the doctor for a medical condition, the doctor suspects three diseases as the cause of the condition. The three diseases are D1, D2, D3, which are marginally independent from each other. There are four symptoms S1, S2, S3, S4 which the doctor wants to check for presence in order to find the most probable cause of the condition. The symptoms are conditionally dependent to the three diseases as follows: S1 depends only on D1, S2 depends on D1 and D2. S3 is depends on D1 and D3, whereas S4 depends only on D3. Assume all random variables are Boolean, they are either `true' or `false'.

a) Draw the Bayesian network for this problem.

b) Write down the expression for the joint probability distribution as a product of conditional probabilities.

c) What is the number of independent parameters that is required to describe this joint distribution?

d) Assume there were no conditional independence between the variables, how many independent parameters would be required then?

Q6 [2M]. You are a witness of a night-time hit-and-run accident involving a taxi in Mumbai. All taxis in Mumbai are blue or green. You swear, under oath, that the taxi was blue. Extensive testing shows that under the dim lighting conditions, discrimination between blue and green is 75% reliable. Assume two random variables: B (taxi was Blue), and LB (taxi Looked Blue). Therefore, P(LB | B) = 0.75, and P (¬LB | ¬B) = 0.75

Is it possible to calculate the most likely color for the taxi? Justify your answer.

Q7 [4M]. Consider the algorithm (Figure-1) for making a CSP arc consistent. Answer the following parts:

- **a)** There are few mistakes in the given AC-3 algorithm. Correct and write the pseudo code again.
- **b)** What is the time complexity of AC-3 algorithm? Assume a CSP with n-variables, each with domain size atmost d, and with c binary constraints.
- **c)** Consider the following map coloring problem (Figure-2). Assign each region one of the three colors (Red, Green, Blue) so that no two adjacent regions have the same color.

```
function AC-3(csp) returns false if an inconsistency is found and true otherwise
  inputs: csp, a binary CSP with components (X, D, C)
  local variables: queue, a queue of arcs, initially all the arcs in csp
  while queue is not empty do
      (Xi, Xj)←REMOVE-FIRST(queue)
      if REVISE (csp, Xi, Xj ) then
          if size of Di = 0 then return false
  return true

function REVISE (csp, Xi, Xj ) returns true iff we revise the domain of Xi
  revised ←false
  for each x in Di do
  if no value y in Dj allows (x, y) to satisfy the constraint between Xi and Xj then
      revised ←true
return revised
```

Figure-1

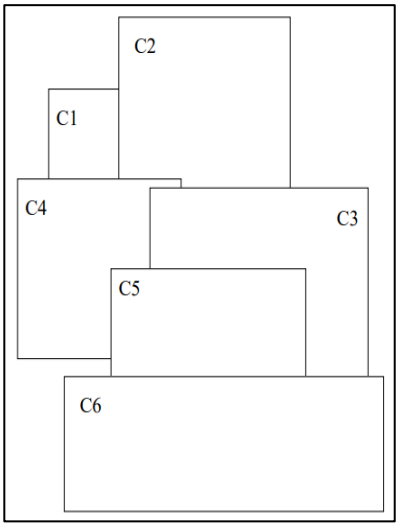

Figure-2

Q8 [5M]. Answer the following:

a) Convert each sentence below to clausal form.

I. $\forall y\ \exists x\ r(x, y) \vee s(x, y)$

II. $\forall y\ (\exists x\ r(x, y)) \rightarrow p(y)$

III. $\forall y\ \exists x\ (r(x, y) \rightarrow p(x))$

b) Consider the vocabulary with only four symbols: A, B, C, and D. For each of the following sentences, how many possible worlds make it True.

I. $(A \wedge B) \vee (C \wedge D)$

II. $\neg (A \wedge B \wedge C \wedge D)$

III. $B \rightarrow (A \wedge B)$

c) Consider the following English statements and corresponding propositions.

1) Either (I like Animals and I also like Trees) or (I like Trees and I like Cats too) [**$(A \wedge T) \vee (T \wedge C)$**]
2) If I like Cats, then I also like Dogs [**$C \rightarrow D$**]
3) If I like Dogs, that implies I like all Animals [**$D \rightarrow A$**]
4) If I like Animals and Trees then I have to be an Environmentalist [**$(A \wedge T) \rightarrow E$**]

Do the resolution proof to answer the question: "*Am I in fact an Environmentalist?*"

Q9 [5M]. Here are two sentences in the language of first-order logic:

(A) $\forall x\ \exists y\ (x \geq y)$

(B) $\exists y\ \forall x\ (x \geq y)$

Assume that the variables range over all the natural numbers 0, 1, 2, . . . , ∞ and that the "≥" predicate means "is greater than or equal to." Under this interpretation,

- **a)** translate (A) and (B) into English.
- **b)** Which of (A) and/or (B) is NOT true under the given interpretation? Justify.
- **c)** Does *(A) logically entail (B)* OR *(B) logically entail (A)*? Justify.
- **d)** Using resolution, try to prove that (A) follows from (B). Do this even if you think that (B) does not logically entail (A); continue until the proof breaks down and you cannot proceed (if it does break down). Show the unifying substitution for each resolution step. If the proof fails, explain exactly where, how, and why it breaks down.
- **e)** Now try to prove that (B) follows from (A).

Q10 [4M]. Answer the following:

a) How to characterize the quality of a heuristic?

b) For eight puzzle problem, consider the following heuristics.
H1: Number of misplaced tiles
H2: Sum of Euclidean distances of the tiles from their goal positions.
With respect to Figure-3, find H1 and H2. Which one these are admissible. Justify.

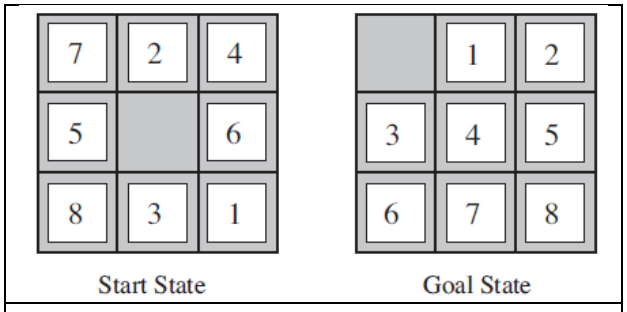

Figure-3

c) Consider the state space shown. It is in the form of a grid of 8 rows and 5 columns. Each square is of 1-unit length and 1-unit breadth. This means an agent moves 1-unit when it moves horizontally (or vertically) from one square to its adjacent one. Also, an agent cannot move in the darkened squares. Apply A* search to move from S (i.e. H1) to G (i.e. A1). Use Euclidean distance as a heuristic. Show the path and all the intermediate steps.