Two hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Fundamentals of Artificial Intelligence

Date: Friday 23rd May 2014
Time: 09:45 - 11:45

Answer One Question from Section A and One Question from Section B.
Use a SEPARATE answerbook for each SECTION.

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

[PTO]
Question 1.

a) In the robot localization problem, why is the probabilistic approach needed? In the probabilistic approach, how is the location of a robot represented? (2 Marks)

b) Let $E$, $F$ and $G$ be three events and $(E \lor F) \land G = \phi$. Use the definition of probability distribution to prove that $p(E \lor F \lor G) = p(E) + p(F \land E^c) + p(G)$, where $E^c$ is the complement of $E$. (3 Marks)

c) Let a single robot be a point object located at some position in a square arena with $2 \times 2$ square grid, in which there is one obstacle occupying position $(0, 0)$. Further, for each position, there are 4 possible orientations $(0, 1, 2, 3)$. Now answer the following questions:

i) Suppose that the robot is equally likely to be in each possible pose. Then what is the initial probability that the robot is located at each pose? (1 Mark)

ii) Based on the initial probability that the robot is located at each pose, give the formulae to calculate the initial probabilities that the robot is located at each position and at each orientation. Then apply the formula to calculate the initial probabilities that the robot is located at position $(1, 1)$ and orientation $1$ respectively (Note: In answering this part of the question you should not assume that the robot is equally likely to be in each possible position or orientation). (4 Marks)

iii) Suppose that the first observation $o$ is obtained from a sensor and the following conditional probabilities are known:

$$p(o \mid L_0, 0, t) = 1/16 \quad p(o \mid L_0, 1, t) = 1/16 \quad p(o \mid L_1, 1, t) = 1/8 \quad (t = 0, 1, 2, 3)$$

Give the formula to calculate the conditional probability that the robot is located at each pose under observation $o$. Then apply the formula to calculate the probabilities that the robot is located at pose $(1, 0, 0)$ and $(1, 1, 1)$ respectively. (5 Marks)

iv) Assume that the robot is located exactly at position $(1,1)$ (that is, the robot is located at this position with probability 1) with the probability of each orientation being given as follows:

$$p(L_{1,1,0}) = p(L_{1,1,1}) = 0.3 \quad p(L_{1,1,2}) = p(L_{1,1,3}) = 0.2$$

Now the robot is asked to take the action $a$ which is to turn right by 1 unit. However, this action cannot be accurately executed and the possible outcomes after action $a$ are turning right by 1, 2, or 3 units. Further it is known that the probabilities to turn right by 1, 2 or 3 units when action $a$ is executed are 0.7, 0.2 and 0.1 respectively. Calculate the probabilities that the robot is located at pose $(1, 1, 3)$ after action $a$ is executed (Note: 4 marks for the step by step process and 1 mark for the correct answer). (5 Marks)
Question 2.

a) What is a Dutch book?  

(1 mark)

b) Let \(a\) and \(b\) be events. Suppose an agent’s degrees of beliefs are as follows:

\[ p(a) = 0.3 \]
\[ p(b) = 0.4 \]
\[ p(a \lor b) = x \]

and the agent will consider the following bets:

<table>
<thead>
<tr>
<th>Bet No</th>
<th>Ticket</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bet 1</td>
<td>£1 if (a); (£0, otherwise)</td>
<td>£0.3</td>
</tr>
<tr>
<td>Bet 2</td>
<td>£1 if (b); (£0, otherwise)</td>
<td>£0.4</td>
</tr>
<tr>
<td>Bet 3</td>
<td>£1 if (a \lor b); (£0, otherwise)</td>
<td>(£x)</td>
</tr>
</tbody>
</table>

Now answer the following questions:

i) If \(x = 0.7\) and \(a \land b \neq \phi\), does the Dutch book occur? Justify your answer. (3 Marks)

ii) If \(p(a \land b) = 0.1\), then what value should be assigned to \(x\) in order to avoid the Dutch book. (1 Mark)

c) In a certain county, 50% of registered voters are Republicans, 40% are Democrats, and 10% are Independents. When those voters were asked about increasing military spending, 40% of Republicans opposed it, 50% of the Democrats opposed it, and 60% of the Independents opposed it. Then what is the probability that a randomly selected voter in this county opposes increased military spending? (6 Marks)

d) Consider the following five-door Monty Hall problem: Step 1, you choose two doors such as doors 1 and 2; Step 2, Monty Hall opens one door with no car behind such as door 3; Step 3. Monty Hall asks you to decide which two doors is your final choice. You will win the car if it is behind either of the two chosen doors. Now apply Bayes’ theorem to calculate the probabilities to win the car by the following three choices: stick, switch one door, and switch both doors. (Note: you need to give detailed formulas and calculations step by step, rather than just present the final result). (9 Marks)
Section B

Answer ONE question

Use a separate answer book for this Section

3. a) It is useful to use a set of features to characterise data used by a classifier. Consider the problems of recognising:

i) an illness from symptoms
ii) isolated spoken words from a restricted vocabulary.

Explain how these problems can be addressed by using a classifier. Describe the types of features that might be useful in each case and explain why. What data would be used to train the classifier and how would you ensure you had a representative sample? (4 marks)

b) Given data described by a vector of features $x$, that can belong to one of two classes, $T$ and $F$, write an expression that relates $p(T|x)$ to $p(x|T)$, $p(x|F)$, $p(T)$ and $p(F)$. Explain the meaning of each term and how the expression can be used to classify data and also give an estimate of the confidence in the classification. (4 marks)

c) A naive Bayes classifier makes the assumption that

$$p(x|T) = \prod_{i=1}^{d} p(x_i|T) \quad \text{and} \quad p(x|F) = \prod_{i=1}^{d} p(x_i|F)$$

where the subscripts of $x$ represent the elements of the feature vector. Consider the logic table shown below:

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_2$</th>
<th>OUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Use the frequency counts from these four examples to estimate $p(x_i|T)$ and $p(x_i|F)$ for $i = 1, 2$, and $p(T)$. Use the naive Bayes assumption to calculate $p(T|x_1 = 1, x_2 = 1)$. How does the classifier perform in this example? (7 marks)

d) Explain how you could use a naive Bayes classifier in a spam email detector. Discuss the features you will use, how you would collect training data and how you would train the classifier. (3 marks)

e) Could this approach be modified to identify the sender of an email from its contents? Why? (2 marks)
4. a) Define a 1\textsuperscript{st} order Markov chain. (2 marks)

b) The following questions refer to the Markov chain model below, which represents the three words “hello”, “hi” and “bye” as a sequence of phonemes.

\[\text{START} \rightarrow \text{hh} \xrightarrow{0.5} \text{eh} \xrightarrow{0.5} \text{l} \xrightarrow{0.5} \text{ow} \xrightarrow{0.5} \text{END}\]

i) The transitions between states are labeled. What do these labels mean? (2 marks)

ii) What is the rule that will let you fill in the missing values for X, Y and Z in the diagram? What are these values? (2 marks)

iii) Explain the role of the self-transitions in the model. (2 marks)

iv) What is the probability of the sequence “hh-hh-ay-ay” according to the model? (2 marks)

v) A sequence of exactly four phonemes is produced by the model. Calculate the probability that this sequence corresponds to the word “hello”. (7 marks)

vi) Explain how the above model could be extended to make a word recognition system. The system would take recorded sound in the form of a WAV file and return the most likely word. What additional processing steps would be required for this to work? How would you set any additional model parameters that may be required, assuming you have access to labelled training data? (3 marks)

END OF EXAMINATION