Fundamentals of Artificial Intelligence

Date: Thursday 27th May 2010
Time: 14.00 – 15.30

Answer One Question from Section A and One Question from Section B.

Use separate Answerbooks 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
Section A

1. a) What is the robot localization problem? List the different sources of uncertainty in the robot localization problem. (3 marks)

b) What were the main activities which led to AI boom from 1980 to 1987? (2 marks)

c) Let a single robot be a point object located at some position in a square arena with 3x3=9 square grids and 10 possible orientations. There is one obstacle occupying 2x2=4 positions in the arena. Suppose we know that the robot is equally likely to be in each possible pose, where a pose is defined as the combination of a position and orientation. Then,

   i) What is the probability that the robot is located at each pose?

   ii) What is the probability that the robot is located at each position?

   iii) What is the probability that the robot is located at each orientation?

   iv) Explain the probability formulas that you use to calculate the probabilities in ii) and iii) above. (5 marks)

d) Let a single robot be a point object located at some position in a square arena with 2x2=4 square grids. The probability of the robot being located at each position \( L_{00}, L_{01}, L_{10} \) or \( L_{11} \) is initially equal. Now a new observation \( o \) is obtained from a sensor. The following conditional probabilities are known:

\[
p(o | L_{00}) = 0 \\
p(o | L_{10}) = 1/4 \\
p(o | L_{01}) = 1/4 \\
p(o | L_{11}) = 1/2
\]

Give the formula to calculate the conditional probability that the robot is located at position \( L_{11} \) under observation \( o \). Apply the formula to calculate the value of this conditional probability. (3 marks)
(Question 1 continues from the previous page)

e) What is the definition of a conditional probability? Based on this definition, prove Bayes' Theorem which is defined as follows: If \( p \) is a probability distribution, and \( E \) and \( F \) are two events satisfying \( p(E) > 0 \) and \( p(F) > 0 \), then

\[
p(E | F) = \frac{p(F | E) p(E)}{p(F)}
\]

(4 marks)

f) Let \( E \) be an event and the agent's degree of belief for the event is \( p(E) \). Then what would the agent consider a fair price for the following ticket and why?

<table>
<thead>
<tr>
<th>Price</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>£x</td>
<td>if ( E )</td>
</tr>
<tr>
<td>£0</td>
<td>if otherwise</td>
</tr>
</tbody>
</table>

(3 marks)
2. a) A widely accepted view or meaning of AI is a computational system that behaves rationally. Using an example, explain why behaving rationally does not always achieve all one’s goals successfully. (2 marks)

b) What is the definition of a probability distribution? Based on the definition, prove that, for any two events $E$ and $F$, if $E \subseteq F$, then $p(E) \leq p(F)$. (5 marks)

c) Let a single robot be a point object located at some position in a square arena with 2x2 square grids. The probability of the robot being located at each position $L_{00}$, $L_{01}$, $L_{10}$, or $L_{11}$ is the same initially. Now the robot takes an accurate action $a$ (such as move forward one unit). The conditional probabilities that the robot is located at the position $L_{11}$ under action $a$ from each position are given as below:

\[
\begin{align*}
p(L_{11} | L_{00} \land a) &= 1/2 \\
p(L_{11} | L_{01} \land a) &= 1/4 \\
p(L_{11} | L_{10} \land a) &= 1/8 \\
p(L_{11} | L_{11} \land a) &= 1/8 
\end{align*}
\]

Which formula can be used to calculate the conditional probability that the robot is located at position $L_{11}$ under action $a$? Apply the formula to calculate the conditional probability. (4 marks)

d) What is a Dutch book? What condition for an agent’s degree of belief needs to be satisfied in order to avoid the Dutch book? Suppose that an agent’s degree of belief for two events $E$ and $F$ are given as follows: $p(E) = 0.5$, $p(F) = 0.4$, $p(E \land F) = 0.3$, and $p[not(E \lor F)] = 0.45$. In this case, will a Dutch book occur and why? (4 marks)

e) Consider the following revised four-door Monty Hall problem: Stage 1, Monty Hall opens a door; Stage 2, you choose a door; Stage 3, Monty Hall opens another door; Stage 4, you stick or switch with your choice. Based on Bayes’ theorem, calculate the probabilities of winning the car when you stick or switch (give detailed formulas and calculations for each step, rather than just presenting the final result). (5 marks)
Section B

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

i) spam email filtering,
ii) spoken word recognition.

Explain how these problems can be addressed by using a classifier. In each case describe a set of features that you think would be useful for training the classifiers? (6 marks)

b) You are given data described by a vector of features $x$. The data can belong to either the class T (true) or the class F (false). Write $p(T | x)$ in terms of $p(x | T)$, $p(x | F)$ and $p(T)$. Explain how this expression can be used to classify data and to associate classifications with a degree of belief. (3 marks)

c) A naïve Bayes classifier makes the simplifying assumption that,

\[
p(x | T) = \prod_{i=1}^{d} p(x_i | T)
\]

\[
p(x | F) = \prod_{i=1}^{d} p(x_i | F)
\]

where $x = [x_1, x_2, ..., x_d]$ is a feature vector. Consider the AND logical operator’s truth table below.

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_2$</th>
<th>AND</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>F</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>F</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>F</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>T</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 naïve 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 would apply a similar naïve Bayes classifier to the spam filtering application discussed in part (a) above. (4 marks)
4. a) Define a 1st 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.

   ![Markov chain diagram]

   i) Explain the role of the self-transitions in the model (shown by the arrows which leave and enter the same state). (2 marks)

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

   iii) What is the probability that a sequence produced by the model will correspond to the word “bye”? (2 marks)

   iv) A sequence of exactly three phonemes is produced by the model. Calculate the probability that this sequence corresponds to the word “bye”. (5 marks)

   v) 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? (7 marks)