On-Line Examination

One and a half hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Artificial Intelligence Fundamentals

Date: Friday 16th May 2008
Time: 14:00 – 15:30

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

The exam will be taken on line.
This is the paper format, which will be available as a backup

The use of electronic calculators is NOT permitted.
Section A

1a. State and prove Bayes' Theorem, relating the quantities \( p(A|B) \), \( p(B|A) \), \( p(A) \) and \( p(B) \), where \( A \) and \( B \) are propositions.

Give carefully any necessary conditions for the theorem you state to be meaningful.

(4 marks)

1b. A mobile robot is equipped with a sensor which measures the distance to the next obstacle in front of it. The sensor returns an integer in the range 0--99 (inclusive). For \( i \geq 2 \) and \( i \leq 97 \), if the true value of the distance to the next obstacle is \( i \), the sensor returns the correct value with 70% probability, has an error of 1 with 20% probability, and an error of 2 with 10% probability. The distribution of errors is symmetric.

Let \( L_i \) be the proposition that the true distance to the next obstacle is \( i \), and let \( O_j \) be the proposition that the sensor returns a reading of \( j \).

Give values for the following. (You may express your answers either as a fraction or a decimal, as you wish.)

(i) \( p(O_{50} | L_{48}) \)

(ii) \( p(O_{50} | L_{49}) \)

(iii) \( p(O_{50} | L_{50}) \)

(iv) \( p(O_{50} | L_{51}) \)

(v) \( p(O_{50} | L_{52}) \)

(vi) \( p(O_{50} | L_i) \) where \( 2 \leq i \leq 47 \) or \( 53 \leq i \leq 97 \)  

(6 marks)

1c. Assuming that the robot knows only that the distance to the next obstacle is in the range 2--97 (inclusive), but otherwise has no idea of its position, give values for the following. (You may express your answers either as a fraction or a decimal, as you wish.)

(i) \( p(O_{50}) \)  

(ii) \( p(L_{50} | O_{50}) \)  

(iii) \( p(L_{51} | O_{50}) \)  

(iv) \( p(L_{52} | O_{50}) \)  

(4 marks)

1d. Suppose the sensor is polled twice (without moving the robot in between), and that the values returned are 50 and 53, respectively. Why would the policy of conditionalization be problematical in this case?

(2 marks)
2a. List three types of on-board sensor which a mobile robot might use to help determine its position. In each case, give the main principle of operation of the sensor.

2b. Give two different sources of uncertainty in the effects of an actuator whose function is to advance a wheeled robot on a flat surface by a given distance.

2c. The following publications or events were historically significant for the development of AI generally, and for the development of the techniques you have used in the laboratory work for this course in particular. What are the dates of these events?

   (i) Bayes' Essay towards Solving a Problem in the Doctrine of Chances, published shortly after the author's death.
       A   B   C   D   E
       1664 1764 1814 1864 1914

   (ii) The conference at Dartmouth College, in the funding application for which the term Artificial Intelligence was used for the first time.
        A   B   C   D   E

   (iii) Alan Turing's paper Computing Machinery and Intelligence
        A   B   C   D   E
        1910 1920 1930 1940 1950

   (iv) The first defeat of Gary Kasparov in a chess game by Deep Blue
        A   B   C   D   E

   (v) First draughts-playing program to run on the Ferranti Mark I computer.
        A   B   C   D   E
        1911 1931 1951 1971 1991

2d. What is the Turing Test? What is it designed to establish? In your opinion, does it serve the purpose for which it was designed? (5 marks)

2e. The first laboratory exercise for this course (on robot localization) ran very slowly when the robot had to take into account the probabilistic effects of actions. In certain situations, however, rapid localization may be required. Suggest a strategy for speeding this process up. (2 marks)

[PTO]
Section B

3a. Explain what the axes represent in a sound waveform. (2 marks)

3b. Explain how a sound waveform can be converted into an uncompressed digital format, e.g. as stored in a WAV file (you may wish to describe the role of sampling and quantization). (3 marks)

3c. A Fourier transform provides an alternative representation of sound. Why is this representation useful in speech recognition? (3 marks)

3d. Give an example of an appropriate application for each of the following speech recognition tasks:

   (i) Isolated word recognition,
   (ii) Continuous speech recognition,
   (iii) Spontaneous speech recognition.

Discuss the relative difficulty of these tasks. For each case, consider whether a language model would be helpful. (6 marks)

3e. The following hidden Markov model (HMM) was used in the labs to distinguish between the words “yes” and “no”. The HMM was trained using sequences of feature vectors extracted from real speech. Each vector in the training sequences was labelled according to the corresponding state: “sil”, “yes” or “no”. Which additional components of the model are not shown? Describe the process of training for these components. (4 marks)

3f. How would you evaluate the performance of the trained HMM from the previous part? (2 marks)
4a. What is a phoneme? Explain why phonemes are useful in speech recognition. (5 marks)

4b. The following questions refer to the Markov chain model below, which represents the words “on” and “off” as a sequence of phonemes.

(i) The numbers associated with arrows leaving each state sum to one. How is this property reflected in the table? Why must this condition hold for a Markov chain model to be valid? (1 mark)

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

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

(iv) A sequence containing exactly two phonemes is generated by the model. Calculate the probability that this sequence corresponds to the word “on”. (6 marks)

(v) What additional model component would be required in order to convert this Markov chain model into a hidden Markov model (HMM)? How does this additional component allow the HMM to model speech data? (4 marks)