

COMPUTER SCIENCE TRIPOS Part IB

Tuesday 7 June 2011 1.30 to 4.30

COMPUTER SCIENCE Paper 4

*Answer **five** questions.**Submit the answers in five **separate** bundles, each with its own cover sheet. On each cover sheet, write the numbers of **all** attempted questions, and circle the number of the question attached.*

<p>You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator</p>

STATIONERY REQUIREMENTS*Script paper**Blue cover sheets**Tags***SPECIAL REQUIREMENTS***Approved calculator permitted*

1 Artificial Intelligence I

A perceptron takes inputs $\mathbf{x}^T = (x_1 \ x_2 \ \cdots \ x_n) \in \mathbb{R}^n$ and computes its output

$$h(\mathbf{x}; \mathbf{w}) = w_0 + \sum_{i=1}^n w_i x_i$$

using weight vector $\mathbf{w}^T = (w_0 \ w_1 \ w_2 \ \cdots \ w_n) \in \mathbb{R}^{n+1}$. We aim to use it to solve a regression problem using a training set $\mathbf{s}^T = ((\mathbf{x}_1, y_1) \ (\mathbf{x}_2, y_2) \ \cdots \ (\mathbf{x}_m, y_m))$ with $y_i \in \mathbb{R}$. The approach will be to minimise the error function

$$E(\mathbf{w}) = \sum_{i=1}^m (y_i - h(\mathbf{x}_i, \mathbf{w}))^2$$

by gradient descent.

(a) Derive the gradient descent learning algorithm for this problem. [5 marks]

(b) The application dictates that the learning process sets as many weights as possible to zero, with the possible side effect that E is increased. It has been suggested that the error function used above might be modified by adding a further term

$$\lambda \sum_{i=0}^n f(w_i, \theta)$$

to E where

$$f(w, \theta) = \begin{cases} 1 & \text{if } |w| > \theta \\ 0 & \text{if } |w| \leq \theta \end{cases}$$

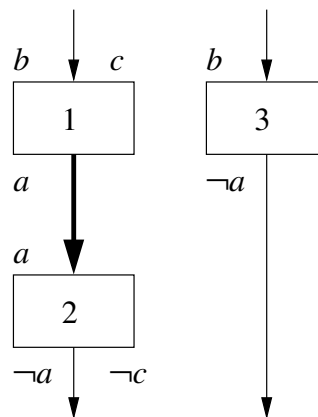
(i) Explain the purpose of the parameters λ and θ in the extra term. [4 marks]

(ii) Assuming we continue to use a gradient descent approach, explain why this term might be inappropriate. [1 mark]

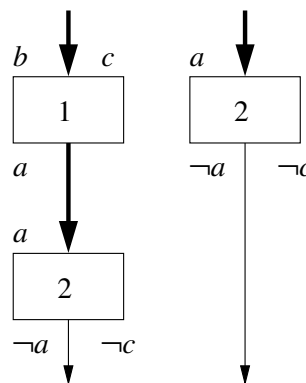
(c) Suggest a function that is appropriate for a gradient descent approach, having a shape similar to that of f , and derive the associated gradient descent learning algorithm. [10 marks]

2 Artificial Intelligence I

- (a) Define what it means for the solution to a partial order planning problem to be *consistent*. [4 marks]
- (b) Define *promotion* and *demotion* in the context of partial order planning, and explain why they are needed. [4 marks]
- (c) The following diagram shows part of an unfinished solution to a partial order planning problem represented in the STRIPS language.



- (i) Explain why this part of the solution is not consistent. [2 marks]
- (ii) Can a single promotion or demotion be used to make this part of the solution consistent? In each case either show how this can be achieved, or give an argument as to why it cannot. [4 marks]
- (d) The following diagram shows a modified part of the unfinished solution, which is still inconsistent. Can a single promotion or demotion be used to make this part of the solution consistent? In each case either show how this can be achieved, or give an argument as to why it cannot. [6 marks]



3 Computer Graphics and Image Processing

You are writing code for a new graphics card that is software programmable, rather than having the algorithms embedded in hardware. You want to write a fast triangle-drawing program to test the card.

- (a) Provide pseudocode, or similar, that draws a triangle with a constant colour. Assume that the inputs are the colour of the triangle and three two-dimensional points, representing the three vertices of the triangle. Further, assume that all three vertices lie on the visible screen and that no anti-aliasing is required. You may assume that there is a function to set a pixel, (x, y) , to a particular colour, c , e.g. `setpixel(x,y,c)`, but you must provide pseudocode for any other functions that you require. Your answer should be sufficiently detailed that it could be transferred directly into a language such as Java but your answer does not, itself, have to be syntactically correct. [13 marks]
- (b) Outline the extra steps required to draw a triangle specified by three-dimensional points in world space, where the triangle may extend beyond the edges of the screen after conversion to screen space. [4 marks]
- (c) Outline the steps required to calculate the triangle's colour, assuming diffuse shading, with multiple point lights, but still producing a single colour for the whole triangle. [3 marks]

4 Computer Graphics and Image Processing

A new computer display is being designed to project large images onto the walls of buildings. It works by reflecting a laser beam off a mirror mounted on a pair of loudspeaker coils. Two analogue signals control the horizontal and vertical deflection of the laser beam by driving the coils. A third digital signal turns the laser on and off. The deflection coils can be driven up to a frequency of 20 kHz, and the laser can be turned on and off at a frequency of 10 MHz. Two approaches are being considered: raster scan and calligraphic.

- (a) What resolution two-level (on/off) raster display could the hardware support? [4 marks]
- (b) How might grey levels be provided? What resolution raster display could the hardware support if eight levels of grey were required? [3 marks]
- (c) An alternative approach would be a random scan calligraphic display that drew lines by deflecting the mirrors to the beginning of a line, turning the laser on, deflecting the mirrors to the end of the line (ensuring that the intermediate deflections lay along a straight line), and turning the laser off again. How many lines could be drawn on this calligraphic display before the image flickered? [2 marks]
- (d) The analogue signals controlling the horizontal and vertical deflection coils are produced by fast digital-to-analogue converters with 10-bit inputs. Present an algorithm for generating intermediate coordinates along the line to ensure that lines drawn on the calligraphic display are straight and also evenly illuminated. [7 marks]
- (e) How could the algorithm be adapted to draw arcs of circles? [2 marks]
- (f) What additional hardware might be useful to assist with generating the intermediate coordinates from the end points of lines? [2 marks]

5 Databases

The lecture notes defined 2nd Normal Form (2NF) as follows. A relational schema R is in 2NF if for every functional dependency $\mathbf{X} \rightarrow A$ either

- $A \in \mathbf{X}$, or
 - \mathbf{X} is a superkey for R , or
 - A is a member of some key, or
 - \mathbf{X} is not a proper subset of any key.
- (a) Present a relational scheme with functional dependencies that is not in 2NF and explain why 2NF is violated. [3 marks]
- (b) Describe an *update anomaly* that could be experienced in a DBMS implementing your schema. [3 marks]
- (c) Decompose your schema into smaller relations that are in 2NF. Justify your answer. [3 marks]
- (d) Is every 2NF schema free from update anomalies? Explain. [3 marks]
- (e) Let us define **Paper 4 Question 5 Normal Form (P4Q5NF)** as follows. A relational schema $R(\mathbf{X})$ with functional dependencies F is in P4Q5NF if for every attribute $A \in \mathbf{X}$ and every key $\mathbf{Y} \subseteq \mathbf{X}$, if $\mathbf{Y} \rightarrow A \in F^+$, then A is prime.

Either prove that P4Q5NF is logically equivalent to 2NF, or provide an example showing that this is not the case. [8 marks]

6 Databases

- (a) Describe NULL values in SQL. [2 marks]
- (b) Discuss **three** problems caused by the introduction of NULL values. [3 marks]
- (c) Can schema normalisation decrease the need for NULL values in a database? Justify your answer. [3 marks]
- (d) What is a *referential integrity constraint* and where are such constraints typically enforced? [3 marks]
- (e) Discuss how schema normalisation can introduce the need for referential integrity constraints. [3 marks]
- (f) Describe the tradeoffs that a database designer needs to consider when deciding whether or not to store data in normalised or de-normalised tables. [6 marks]

7 Economics and Law

- (a) List **five** types of auction. [5 marks]
- (b) Which of the auction types you list in part (a) are strategically equivalent, and why? [5 marks]
- (c) Describe the “winner’s curse”. [5 marks]
- (d) To which type are Ebay auctions strategically equivalent? Justify your answer. [5 marks]

8 Security I

- (a) In Windows NTFS, each file can have an associated access control list (ACL). Each entry has a type in $\{allow, deny\} \times \{explicit, inherited\}$.
- (i) What restriction does the Windows Explorer graphical user interface impose on the order in which these types of access-control entries can appear in an ACL? [4 marks]
- (ii) Give **one** example of a POSIX file access-control configuration for which an equivalent NTFS ACL violates this GUI restriction. [4 marks]
- (b) Your colleagues used a pseudo-random function $f : \{0, 1\}^{64} \rightarrow \{0, 1\}^{64}$ in order to construct a permutation $g : \{0, 1\}^{192} \rightarrow \{0, 1\}^{192}$. The argument and return values of g are split into three 64-bit registers, respectively: $g(X_1, X_2, X_3) = (Y_1, Y_2, Y_3)$. The output of g is calculated as $Y_2 = f(X_1) \oplus X_2 \oplus f(X_3)$, $Y_1 = X_1 \oplus f(Y_2)$, and $Y_3 = X_3 \oplus f(Y_2)$, where \oplus denotes bit-wise exclusive or.
- (i) Show that g is indeed a permutation. [4 marks]
- (ii) Show how an attacker who does not know f can efficiently distinguish g from most random permutations, after evaluating g on two different inputs. [4 marks]
- (iii) After you point out this shortcoming to your colleagues, they propose an improved variant $g'(X_1, X_2, X_3) = (Z_1, Z_2, Z_3)$ that adds another round to g : $Z_1 = Y_1$, $Z_2 = f(Y_1) \oplus Y_2 \oplus f(Y_3)$, and $Z_3 = Y_3$.

Show how this variant still does not fix the problem of efficient distinguishability from most random permutations. [4 marks]

END OF PAPER