COMPUTER SCIENCE TRIPOS Part II

Tuesday 5 June 2001 1.30 to 4.30

Paper 7

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.*
Write on **one** *side of the paper only.*

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

1 Specification and Verification I

(*a*) What is the difference between partial and total correctness? [4 marks]

(*b*) Why is the assignment axiom more problematical for total correctness than for partial correctness? [4 marks]

(*c*) State the `WHILE`-rule for total correctness. [4 marks]

(*d*) What needs to be added to the method of verification conditions to make it work for total correctness? [4 marks]

(*e*) Explain how a total correctness specification $[P]C[Q]$ can be embedded as a term $\mathsf{TotalSpec}\ p\ c\ q$ in higher-order logic. [4 marks]

2 Specification and Verification II

(*a*) Describe the difference between the simple and threshold switching models of transistors. [4 marks]

(*b*) Give and explain an example of a circuit that illustrates when the threshold switching model is superior to the simple switching model. [4 marks]

(*c*) Describe how combinational devices can be modelled as zero-delay sequential devices. When is this appropriate? [4 marks]

(*d*) Write down a CTL formula expressing the property: `Ack` *is true on all paths sometime between 2 units and 5 units of time later*. [4 marks]

(*e*) Describe how an edge-triggered D-type register can be abstracted to a unit delay. Give the formula in higher-order logic that expresses the abstraction. [4 marks]

3 Comparative Architectures

A processor is to be selected for use in a battery powered embedded device that will incorporate the functions of a mobile phone, Personal Digital Assistant (PDA), web browser, and audio/video player.

(*a*) Discuss how the above requirements might influence the instruction set architecture, micro-architecture, and physical implementation of a processor suitable for the above device. [15 marks]

(*b*) Describe how you would go about determining whether a candidate CPU would provide sufficient performance to implement fully the embedded device's specification. [5 marks]

4 Optimising Compilers

An expression is *very busy* at a program point n if, no matter what path is taken from n, some occurrence of the expression is always evaluated before any of the variables appearing in it are redefined. A transformation using *Very Busy Expression* (VBE) analysis is to evaluate the expression at n and store its value for later use.

(*a*) Give dataflow equations for, and a pseudo-code algorithm to calculate, VBE for a program in flowgraph form. State whether your dataflow equations are *forwards* or *backwards*. Sketch the above transformation which exploits VBE in more detail. [11 marks]

(*b*) Show how to calculate the *call graph* of a program, and explain the safety property your call graph should have. (You should relate the call graph you define to possible execution behaviour.) Detail how you handle procedure-valued variables, and state whether it is possible to improve on the technique you have chosen for such variables. [6 marks]

(*c*) Argue how feasible it is to calculate the call graph for a Java program, considering carefully the case of inheritance and use of the `final` keyword. [3 marks]

5 E-Commerce

One of your customers complains about the delivery of goods that she claims she did not order.

(*a*) What is needed to establish that a contract actually existed? [4 marks]

(*b*) What other information is it useful to keep about your customers, and what could you use it for? [4 marks]

(*c*) Describe some of the problems of anonymous payment schemes such as digital cash. Where might they be useful? [8 marks]

(*d*) What recourse would the customer have if she had paid using such a system, and how would you detect whether such a complaint was fraudulent? [4 marks]

6 Security

You are developing a multi-user computer game, and wish to make it harder for players to cheat.

(*a*) Discuss the possible benefits of using

(*i*) encryption/authentication [4 marks]

(*ii*) virus detection technology [4 marks]

(*iii*) intrusion detection techniques [4 marks]

(*b*) What might be the advantages and disadvantages of issuing players with a smartcard and reader? [8 marks]

7 Neural Computing

(*a*) (*i*) In a Hopfield neural network configured as an associative memory, with all of its weights trained and fixed, what *three* possible behaviours may occur over time in configuration space as the net continues to iterate in response to a given input? [3 marks]

(*ii*) How many stable content-addressable memories would you expect a fully connected Hopfield network consisting of 100 neurons to be capable of storing? [1 mark]

(*iii*) What property of those memory patterns would make it most probable that you could successfully train the network to store the maximum number, and why? [3 marks]

(*b*) Explain how five independent dimensions of visual processing are multiplexed together into the three available spatial dimensions of neural tissue, by the structure of the cubic millimetre hypercolumns in the brain's visual cortex. [5 marks]

(*c*) The retina is often regarded as an image capture device; but it has about 100 million input sensors (photoreceptors) yet only 1 million output fibres (optic nerve axons). What are some implications of this 100-to-1 ratio of input channels to output channels? [4 marks]

(*d*) Provide some quantitative evidence and arguments supporting the proposition that: "Connectivity is the basic computational principle in the brain". [4 marks]

8 Computer Systems Modelling

(*a*) What criteria would you consider when selecting between a model based on queueing theory and one based on simulation? When might you use both approaches? [5 marks]

(*b*) Describe the structure of a *discrete event simulator*. What is the principal data structure involved? [5 marks]

(*c*) A queueing network is characterised by a set of *visit counts*, V_i, and *per-visit service requirements*, S_i, for each of N devices. Derive upper bounds on the system throughput (*i*) when the load is very low and (*ii*) as the load tends to infinity. [5 marks]

(*d*) In what situations may the bounds be particularly imprecise? What can be done to construct tighter bounds for the system throughput? [5 marks]

9 Advanced Graphics and HCI

(*a*) The marching squares algorithm is a two-dimensional version of marching cubes – where you generate line segments in 2D rather than triangles in 3D. It could be used, for example, where you have a regular grid of height values and want to draw contours of constant height. Sketch an implementation of this two-dimensional marching squares algorithm. [6 marks]

(*b*) (*i*) Show how to find the first intersection between a ray and a finite-length, open-ended cone, centred at the origin, aligned along the x-axis, for which both ends of the finite-length are on the positive x-axis (i.e. $0 < x_{min} < x_{max}$). [6 marks]

(*ii*) Extend this to cope with a closed cone (i.e. the same cone, but with end caps). Take care to consider any special cases. [5 marks]

(*iii*) Extend this further to give the normal vector at the intersection point. [3 marks]

10 Advanced Graphics and HCI

(*a*) Name *two* interaction devices or techniques already in widespread use at the time the mouse was developed. How was the mouse related to *each* of them? Describe *two* ways in which the mouse has influenced usability evaluation theories. [6 marks]

(*b*) Using *two* Cognitive Dimensions of Notations as examples, describe how *each* could influence some design decision involved in the user interface of a presentation software product (Microsoft PowerPoint is an example of presentation software). [4 marks]

(*c*) Describe a simple model of human cognition, at the level of functions such as "long term memory". How do the evaluation methods of KLM, GOMS, Cognitive Walkthrough, and Heuristic Evaluation relate to this model? [10 marks]

11 Natural Language Processing

Define *two* of the following four types of ambiguity, giving examples. For each of these two types, describe *one* possible technique for resolving such ambiguities.

(a) Anaphoric coreference ambiguity.

(b) Speech act identification ambiguity.

(c) Part-of-speech assignment ambiguity.

(d) Prepositional phrase attachment ambiguity.

[10 marks each]

12 Information Theory and Coding

(a) Consider n different discrete random variables, named $X_1, X_2, \ldots, X_n$, each of which has entropy $H(X_i)$.

Suppose that random variable X_j has the smallest entropy, and that random variable X_k has the largest entropy.

(i) What is the upper bound on the joint entropy $H(X_1, X_2, \ldots, X_n)$ of all these random variables? [2 marks]

(ii) Under what condition will this upper bound be reached? [2 marks]

(iii) What is the lower bound on the joint entropy $H(X_1, X_2, \ldots, X_n)$ of all these random variables? [2 marks]

(iv) Under what condition will this lower bound be reached? [2 marks]

(b) (i) Define the Kolmogorov algorithmic complexity K of a string of data. [2 marks]

(ii) What relationship is to be expected between the Kolmogorov complexity K and the Shannon entropy H for a given set of data? [2 marks]

(iii) Give a reasonable estimate of the Kolmogorov complexity K of a fractal, and explain why it is reasonable. [2 marks]

(c) The signal-to-noise ratio SNR of a continuous communication channel might be different in different parts of its frequency range. For example, the noise might be predominantly high frequency hiss, or low frequency rumble. Explain how the information capacity C of a noisy continuous communication channel, whose available bandwidth spans from frequency ω_1 to ω_2, may be defined in terms of its signal-to-noise ratio as a function of frequency, $SNR(\omega)$. Define the bit rate for such a channel's information capacity, C, in bits/second, in terms of the $SNR(\omega)$ function of frequency.

[Note: This question asks you to generalise beyond the material lectured.]

[6 marks]

13 Types

(*a*) What does it mean to say that an ML type scheme σ *generalises* an ML type τ? Writing $\sigma \succ \tau$ for this relation and defining

$$\sigma_1 = \forall\{\alpha,\beta\}(\alpha \to \beta) \qquad \sigma_2 = \forall\{\alpha\}(\alpha \to \beta)$$
$$\tau_1 = (\alpha \to \beta) \to \alpha \qquad \tau_2 = (\beta \to \alpha) \to \beta$$

say whether or not $\sigma_i \succ \tau_j$ holds for each of the four possibilities. [5 marks]

(*b*) Give the axioms and rules for inductively generating ML typing judgements of the form

$$\Gamma \vdash M : \tau$$

where $\Gamma = (\Gamma_{\mathrm{tv}}, \Gamma_{\mathrm{ta}})$ with Γ_{tv} a finite set of type variables, Γ_{ta} a finite function mapping some variables to type schemes whose free type variables are in Γ_{tv}, and τ is a type whose type variables are in Γ_{tv}. You may restrict attention to expressions M involving variables, boolean values, conditionals, function abstraction and application, and let-expressions. [8 marks]

(*c*) Explain why the expressions $\mathbf{let}\, x = M \,\mathbf{in}\, M'$ and $(\lambda\, x\, (M'))M$ can have different typing properties in the ML type system even though their evaluation behaviour is the same; illustrate your answer by taking M to be $\lambda\, y\, (y)$ and M' to be $x\, x$. [7 marks]

14 Additional Topics

(*a*) Briefly describe a technology which can generate containment location information (e.g. room). [4 marks]

(*b*) Briefly describe a technology which can generate co-ordinate location information (e.g. latitude/longitude). [4 marks]

(*c*) For each technology state how accurate the location information is likely to be both in terms of precision and erroneous readings. [6 marks]

(*d*) For each technology give *one* application and discuss how the precision affects the end-user. [6 marks]

15 Additional Topics

(*a*) Describe the major functions of the Space, Control and User segments of the Navstar Global Positioning System. [8 marks]

(*b*) Explain the meaning of the term *pseudo-range*, and why at least four pseudo-ranges are necessary for a three-dimensional position fix. [4 marks]

(*c*) How does the geometry of the satellites affect the accuracy of the position fix, particularly in built-up areas? Illustrate your answer with a sketch. [4 marks]

(*d*) If an alternative satellite system measured ranges directly, perhaps by round-trip timing, how might the geometrical effects in built-up areas be altered? Illustrate your answer with a sketch. [4 marks]

END OF PAPER