COMPUTER SCIENCE TRIPOS Part II

Thursday 2 June 1994 1.30 to 4.30

Paper 9

Answer **five** *questions.*
Submit the answers in five **separate** *bundles each with its own cover sheet.*
Write on **one** *side of the paper only.*

1 Comparative Architectures

Discuss alternative ways of implementing 8-bit and 16-bit object access on processors with a 32-bit or 64-bit bus size. [7 marks]

How might one deal with 9-bit objects? [3 marks]

Both the DEC Alpha and the MIPS R2000 claim to be RISC machines. One has delayed branches and loads but the other does not. Sketch possible design or implementation differences. [7 marks]

Why might *conditional move* instructions be useful? [3 marks]

2 Developments in Technology

Either A multimode graded index fibre has a core of radius a and refractive index $n(r)$, where r is the radius measured from the fibre axis. Show that the equation of a light ray launched at radius r_0 along the fibre core, parallel to the fibre axis, is given by

$$\frac{d^2r}{dz^2} = \frac{1}{2\beta^2}.\frac{dn^2}{dr}$$

where $\beta = n(r_0)$. [8 marks]

The fibre core has a parabolic variation of index with radius, described by

$$n^2(r) = n_{co}^2(1 - \alpha r^2)$$

Show that a short length of the fibre can act as a lens, and derive an expression for its focal length in terms of α, n_{co}, and β. [8 marks]

Explain qualitatively how such a fibre minimises the spreading of an optical pulse propagating along it. [4 marks]

or Describe any one optical system in a compact disc player. In your answer pay particular attention to the following:

- how information is recorded and read from the disc's surface
- the limitations on the disc's storage capacity
- the laser focusing system
- the optical tracking system [20 marks]

3 Distributed Systems

A certain university proposed a design for a workstation with hardware-enforced protection of the system software. Trusted software could then be run on the workstation.

Discuss the advisability of this approach to workstation design. [20 marks]

4 Computer Systems Modelling

A distributed database is organised as four servers with 100 terminals attached via 10 terminal concentrators. The concentrators and the servers are connected by a 100 Mbps network. Each server has a single disk. A typical query generates the following load on the system:

terminal concentrator	70 ms
server CPU	60 ms
disk	25 ms
network	1 Kbyte transfer

Give estimates and bounds of average system response time when each terminal generates

(a) a query every minute

(b) a query every 10 seconds

(c) a query every 5 seconds

[Recall that in a balanced system the utilisation of each device is given by

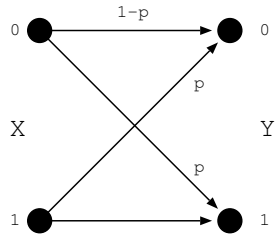
$$U = \frac{N}{N + K - 1}$$

where N is the number of customers and K is the number of devices.]

[20 marks]

5 Information Theory and Coding

Define the mutual information and channel capacity of a discrete memoryless channel. [10 marks]



For the binary symmetric channel with transition probability p, derive the channel capacity. Sketch a graph of the channel capacity against the transition probability. [10 marks]

6 Designing Interactive Applications

Define in one sentence the rôle of requirements in system design. Explain the difference in purpose between performance requirements and functional requirements for interactive systems. [2 marks]

The rules enforced in the City for dealings in equities require that deals (buying and selling stock) are to be recorded within 60 seconds of agreeing each deal. In one particular bank this is currently done by having the dealer write the details of the deal on a paper "ticket" and place it in a pot for collection by a deal-input clerk who enters the deal into an on-line system. Six items of information must be written on the ticket (average lengths shown in parentheses):

1. the customer's account number (four digits)
2. the name of the buyer/seller (six characters)
3. the name of the stock (six characters)
4. the number of shares bought or sold (four digits)
5. the price agreed (five characters; for example, 72 3/8 or 167.5)
6. whether this is a buy or a sell

Each dealer handles only a limited range of stocks, at most thirty (i.e., thirty companies such as ICI, Wellcome, Hanson, etc.). During busy periods, dealers may make deals every 30 seconds.

It is proposed that, in the future, dealers input their deals themselves.

(a) Using the Keystroke-level model, calculate how long it would take dealers to record deals by typing the details on an alphanumeric keyboard. [4 marks]

(b) Suppose all of the names of stocks were displayed on the dealer's interactive workstation, together with their current prices. Sketch out a design to allow the dealer to input some of the details of the deal graphically with a mouse and some with the keyboard. Again using the Keystroke-level model, calculate the speed of entry of deals. Using this result and the result from (a), write a performance requirement for a system to support dealers in entering deals. [8 marks]

Use the following times in seconds for operators: **K** (key-press) = 0.30, **P** (point at target) = 1.10, **H** (home hands to another device) = 0.40, **M** (mentally prepare) = 1.35.

(c) Discuss the benefits of these two alternatives to the current method, and the feasibility of meeting the performance requirements derived in (b) by using other styles of interaction. [6 marks]

7 Algebraic Manipulation

Explain how polynomials can be represented within an algebra system using (a) recursive, and (b) distributed data structures.

For each, illustrate your answer by showing what structure will be generated for the polynomial $a_1 a_2 \ldots a_{10}(1 + x + \cdots + x^{10})$ first when the 11 variables $a_1, \ldots a_{10}$ and x are sorted in alphabetic order, and then when they are kept in inverse alphabetic order. [16 marks]

Does the ordering used make any difference to the amount of space consumed? [4 marks]

8 Artificial Intelligence II

Explain how genetic algorithms differ from conventional mathematical methods for optimisation. [10 marks]

What are the advantages and disadvantages of genetic algorithms? [10 marks]

9 Database Topics

The Relational Model of Data and its associated Data Manipulation Languages contained no provision for handling sets of tuples. Discuss the extent to which the SQL GROUP BY clause overcomes this weakness. [4 marks]

The Nested Relational Model attacks the problem by extending the Data Definition Language that specifies a relational schema. Describe the extension, and explain with the aid of examples how it enables both sets and compound values to be represented. [7 marks]

Explain how aggregates are handled in the Functional Data Model and the Data Language DAPLEX. [6 marks]

In what way can the purging of duplicates cause problems when evaluating functions defined over sets? How does DAPLEX allow the programmer to retain duplicates when they are needed? [3 marks]

10 Numerical Analysis II

Apply the Newton-Raphson formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

for solving the equation $f(x) = 0$ using $f(x) = x - \frac{1}{2}e^{1-x^2}$. Simplify your formula so that the exponential function need be called only once per iteration. [4 marks]

This function has a zero at approximately 0.76. Use $x_0 = 1$ and perform one Newton-Raphson iteration to calculate x_1 and verify that this is a good starting value. Now use $x_0 = -1/\sqrt{2}$ and re-calculate x_1. (You may assume for this purpose: $\sqrt{2} \simeq 1.4,\ 1/\sqrt{2} \simeq 0.71, \sqrt{e} \simeq 1.6,\ 1/\sqrt{e} \simeq 0.61$.) [3 marks]

By sketching the graph of $f(x)$, or otherwise, explain these results. [4 marks]

Perform two Newton-Raphson iterations (in exact arithmetic) for the function $f(x) = x - 1/(x+1)$, with $x_0 = 1$. Suppose that the second iteration is modified as follows:

$$\tilde{x}_2 = x_1 - \frac{f(x_1)}{f'(x_0)}.$$

Examine $|x_2 - \tilde{x}_2|/|x_2|$ to estimate the relative loss of accuracy if the modified method is used. [5 marks]

Although $\tilde{x}_2$ is obviously less accurate than x_2 in general, explain briefly why this modified Newton method is more useful for an n-dimensional problem. [4 marks]

11 Specification and Verification of Hardware

Discuss the problems of providing tractable models of transistors suitable for hardware verification by formal proof. Compare and contrast at least two different models. Illustrate your discussion with concrete examples of transistor circuits.
[20 marks]

12 Semantics of Programming Languages

State the Tarski-Knaster fixed-point theorem. Give a brief justification for the importance of the fixed-point theorem in denotational semantics. [8 marks]

Prove that the least fixed-point operator fix is a continuous function. You may assume the following result:

> Let $\langle D, \sqsubseteq \rangle$ be a complete partial order. Every doubly-increasing chain $\langle d_{ij} \rangle_{i,j \in \omega}$ in D (i.e. for any i, j, i', j' in ω, if $i \leqslant i'$ and $j \leqslant j'$ then $d_{ij} \sqsubseteq d_{i'j'}$) has a least upper bound l. Further,
>
> $$l \;=\; \bigsqcup_{i \in \omega} \bigsqcup_{j \in \omega} d_{ij} \;=\; \bigsqcup_{j \in \omega} \bigsqcup_{i \in \omega} d_{ij} \;=\; \bigsqcup_{k \in \omega} d_{kk}$$

[12 marks]

13 Types

Describe the types and terms of the second-order lambda calculus ($\lambda 2$) and define the type assignment relation for $\lambda 2$. [5 marks]

Are the following expressions typeable in $\lambda 2$? Justify your answer in each case.

$$\begin{aligned} L &= \Lambda\alpha.\Lambda\beta.\lambda x : \alpha.\Lambda\gamma.\lambda f : \alpha \to \gamma.\lambda g : \beta \to \gamma.fx \\ R &= \Lambda\alpha.\Lambda\beta.\lambda y : \beta.\Lambda\gamma.\lambda f : \alpha \to \gamma.\lambda g : \beta \to \gamma.gy \\ S &= \Lambda\alpha.\lambda x : \alpha.x(x_\alpha) \end{aligned}$$

[5 marks]

Find a type σ that makes the following expression typeable in $\lambda 2$ and give, with justification, the type of C.

$$C \;=\; \Lambda\alpha.\Lambda\beta.\Lambda\gamma.\lambda f : \alpha \to \gamma.\lambda g : \beta \to \gamma.\lambda z : \sigma.(z_\gamma f)g$$

[5 marks]

Explain what is meant by *β-reduction* and *β-normal form* for $\lambda 2$ terms. Calculate the β-normal form of the term $C_{\alpha\beta\gamma} f g (L_{\alpha\beta} M)$ where C and L are as above, α, β, γ are type variables, f, g are identifiers, and M is a term in β-normal form with no free identifiers or free type variables. [5 marks]

14 Computational Number Theory

Describe the elliptic curve method for factorising integers. [7 marks]

Derive an estimate, using a heuristic argument, for the expected running time of this method. [13 marks]