COMPUTER SCIENCE TRIPOS Part II

Thursday 8 June 1995 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 VLSI

Summarise the following styles of chip design:

(a) gate array [4 marks]

(b) standard cell [4 marks]

(c) full custom [4 marks]

In which way are these techniques used for implementation of state-of-the-art microprocessors? [8 marks]

2 Information Theory and Coding

Provide the formula for the Discrete Fourier Transform. [4 marks]

Show how it is possible to split the transform of $2M$ points into two transforms of M points. Hence derive the number of complex multiplications and divisions required for the Fast Fourier Transform of 2^N points. [12 marks]

Explain the memory requirements of the Fast Fourier Transform. [4 marks]

3 Digital Communication II

Describe the ISO OSI reference model, including details of the major functions at each level. [14 marks]

At which level is end-to-end encryption normally considered to reside in this model? [2 marks]

Discuss circumstances in which encryption might be used at other protocol levels and the attacks against which such use is defending. [4 marks]

4 Concurrency Theory

Define the notions of *strong equivalence* ($\sim$) and *observation equivalence* ($\approx$) for CCS agents. [4 marks]

A CCS agent P is called τ*-free* if τ does not occur in P, or in the definitions of any constants occurring in P. For any τ-free agent P, show that $\tau.P$ is strongly equivalent to a τ-free agent. (You may assume that there is at least one name, a, which does not occur in the syntactic sort of P.) [6 marks]

A CCS agent expression is called *normal* if it is of the form $\sum_{i \in I} \ell_i.P_i$, where all the ℓ_i are labels (that is, not τ actions). Show that the property

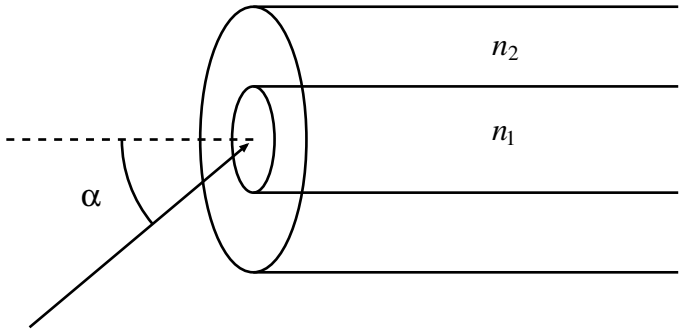
$$(*)\ P_1 \approx P_2 \text{ and } Q_1 \approx Q_2 \text{ implies } P_1 + Q_1 \approx P_2 + Q_2$$

holds for all normal agents P_1, P_2, Q_1, and Q_2. [5 marks]

Does $(*)$ hold for all τ-free agents P_1, P_2, Q_1, and Q_2? [5 marks]

5 Developments in Technology

Light is incident from air on the end face of a multimode optical fibre at angle of incidence α as shown below.



The refractive indices of the core and cladding are n_1 and n_2 respectively, where $\mid n_1 - n_2 \mid \ll 1$. Prove the following condition for the incident light to be guided by the fibre

$$\sin \alpha \leqslant \sqrt{n_1^2 - n_2^2}$$

[4 marks]

What is the main cause of pulse spreading in a step index multimode fibre? [3 marks]

Show, for the same fibre parameters as given above, that the bandwidth $\times$ length product for the fibre is given approximately by

$$B.L = \frac{nc}{(n_1^2 - n_2^2)}$$

where c is the speed of light in free space, and the approximation $n_1 \approx n_2 \approx n$ has been made. [5 marks]

Explain carefully how, by appropriate design of the multimode fibre, the bandwidth might be increased. [3 marks]

Explain carefully what is meant by the term *material dispersion* in an optical fibre, explaining its relative importance for single and multimode fibres. How might it be minimised? [5 marks]

6 Security

The Tatebayashi–Matsuzaki–Newmann protocol may be described as follows:

$$\begin{array}{lcll} A & \rightarrow & S & : \quad {r_A}^3 \pmod N \\ B & \rightarrow & S & : \quad {r_B}^3 \pmod N \\ S & \rightarrow & A & : \quad r_A \oplus r_B \end{array}$$

Explain what is happening here, including the goal and the assumptions. [4 marks]

How can this protocol be attacked? [10 marks]

It has been suggested that the protocol can be strengthened by changing the contents of the last message to $\{r_B\}_{r_A}$ (r_B encrypted by r_A using a secret key algorithm). Does this help? Explain your answer. [6 marks]

7 Optimising Compilers

Carefully define what it means for a function to be *strict* in its i^{th} argument. [4 marks]

Carefully describe how a safe approximation to the strictness properties of a mutually recursive set of functions can be calculated, illustrating your method using the following definitions:

```
f(x,y,z) = h(g(x,y), g(y,z))

g(a,b) = if h(a,b)=0 then a else f(a-1,b,a)

h(p,q) = if p=0 then f(q,p-1,p) else if q=0 then 1 else 0
```

[10 marks]

Discuss how strictness information can be used in the optimisation of pure functional languages on

(a) simple single processor machines

(b) parallel processing hardware

[6 marks]

8 Prolog for Artificial Intelligence

Consider the task of normalising sum expressions. For example, the sums $(a + b) + (c + d)$ and $(a + (b + (c + d)))$ may be normalised into a standard form that is left associative: $a+b+c+d$ or equivalently $((a+b)+c)+d$. Write a Prolog procedure to define predicate `normsum` such that the goal `normsum(X,Y)` succeeds when the sum expression `X` normalises to `Y`. Procedures not using the technique of difference structures will not receive full marks. [20 marks]

9 Running a Business

Explain the difference between a *profit and loss account* and a *cash-flow statement.* Under what circumstances would they show the same figures? [5 marks]

A small software company is offered a development contract, valued at £100,000 (excluding VAT), with 10% to be paid at the start of the contract, 30% invoiced at the first milestone (estimated after 3 months), 50% invoiced on completion, with 10% to be retained for 3 months after completion, as a guarantee against errors.

The company estimates that the project will require 6 months' work from each of two staff, whose annual salary costs are £36,000 and £24,000 respectively. Other overheads are approximately 120% of salary costs.

Draw up an outline monthly profit and loss account and cash-flow statement for the project, ignoring VAT and bank interest. Salary and overheads are charged to the project only while the programmers are actually working on it. [5 marks]

What is the eventual profit the company expects to make, if it undertakes the project, and how much working capital will the project require? [5 marks]

The effort in the project turns out to be underestimated, and the company delivers the first milestone 1 month late, and completes 2 months late compared with the original schedule, requiring both programmers to work for the extra 2 months. How has this affected the profitability and working capital requirement? [5 marks]

10 Natural Language Processing

Imagine you were to construct a natural language processing system for solving syllogisms expressed in English, behaving approximately as indicated:

User:	All men are mortal.	System:	OK
User:	Socrates is a man.	System:	OK
User:	Is Socrates mortal?	System:	Yes

Describe the components of such a system and how they would fit together.

[20 marks]

11 Computer Systems Modelling

In queueing networks, what is meant by a *closed* system? [4 marks]

Consider two closed systems. One has two devices, A and B, and three customers, the other three devices, A, B and C, and two customers. Both have exponentially distributed service times which are device dependent but customer independent. In the first system a customer completing service at a device always moves to the other device. In the second system a customer completing service moves to one of the other two devices with equal probability.

Draw state diagrams for the Markov chains representing these systems. Choose one system to solve for device utilisation in terms of service rates. [10 marks]

For the chosen system, when the service rates are equal does the utilisation of each device correspond to that for a balanced system ($U = \frac{N}{N+K-1}$ where N is the number of customers and K the number of devices)? [3 marks]

Describe the state space for a Markov chain for one of the systems if the service rates were both customer and server dependent. [3 marks]

12 Types

Consider the following datatype and function declarations in Standard ML:

```
datatype tree = Leaf | Node of tree * tree ;
fun iter x f Leaf = x
  | iter x f (Node(y,z)) = f(iter x f y)(iter x f z) ;
```

You are required to encode the datatype `tree` as a closed type τ in the second-order lambda calculus, $\lambda 2$. Find a suitable type τ and closed $\lambda 2$ terms in β-normal form, L, N, and I say, corresponding to `Leaf`, `Node` and `iter` respectively. You should demonstrate for your choices that

$$\begin{aligned} &\vdash L : \tau \\ &\vdash N : \tau \to \tau \to \tau \\ &\vdash I : \forall\alpha.\alpha \to (\alpha \to \alpha \to \alpha) \to \tau \to \alpha \end{aligned}$$

are derivable typing assertions, and that $I_\alpha x f L$ and $I_\alpha x f (N y z)$ are β-convertible to the $\lambda 2$ terms corresponding respectively to the right-hand sides of the clauses in the declaration of `iter`. [14 marks]

Now add to the above Standard ML declarations the function declarations

```
fun rev Leaf = Leaf
  | rev (Node(y,z)) = Node(rev z, rev y) ;
fun div Leaf = Leaf
  | div (Node(y,z)) = div(Node(z,y)) ;
```

Using I, or otherwise, show that there is a closed $\lambda 2$ term of type $\tau \to \tau$, R say, for which RL and $R(Nyz)$ are β-convertible to the $\lambda 2$ terms corresponding respectively to the right-hand sides of the clauses in the declaration of `rev`. Is there a closed $\lambda 2$ term D with similar properties for the declaration of `div`? [6 marks]

13 Complexity Theory

It turns out that the following program, when run using floating-point arithmetic that remains accurate to N decimal places, will compute and print the value of π correct to almost N places. The loop (which has as its main part a step which replaces the values in `a` and `b` by their arithmetic and geometric means, respectively) will be traversed about $\log(N)$ times.

```
a := 1;
b := 1/sqrt(2);
u := 1/4;
x := 1;
pn := 4;
do { p := pn;
     y := a; a := (a+b)/2; b := sqrt(y*b);
     u := u-x*(a-y)*(a-y);
     x := 2*x;
     pn := a**2/u; } while (pn<p);
print(p);
```

You are provided with procedures that can compute Fourier Transforms and their inverses with a transform on k points (using floating-point arithmetic), taking time proportional to $k \log k$. Explain how you could implement the high-precision arithmetic needed to make this program run fast. Do not discuss how the Fourier transform will be implemented — just how it is used, and assume that the floating-point accuracy achieved by the transform will be adequate for your purposes.

[14 marks]

Overall how long (as a function of N) would you expect the complete program to take to run? [6 marks]

You do not need to understand how or why this particular calculation arrives at a value for π, or why the loop is executed only $\log(N)$ times.

14 Semantics

The language IMP′ comprises integer and boolean expressions and commands, defined by

$$\begin{array}{rcl} ie & \in & Iexp \ ::= \underline{n} \mid x \mid ie_1 \ \underline{iop} \ ie_2 \\ be & \in & Bexp \ ::= \underline{b} \mid ie_1 \ \underline{bop} \ ie_2 \\ C & \in & Com \ ::= \texttt{skip} \mid x := \ ie \mid (C_1; C_2) \mid \\ & & \qquad \texttt{if } be \texttt{ then } C_1 \texttt{ else } C_2 \mid \texttt{repeat } C \texttt{ until } be \end{array}$$

where $n \in \mathbb{Z}$, $b \in \{true, false\}$, $iop \in \{+, \times, -\}$, $bop \in \{<, =\}$ and $x \in Pvar$, a set of program variables.

(a) Give an annotated evaluation semantics for IMP′, expressing the usual behaviour of each command and expression form, which derives statements of the forms

$$ie, S \Rightarrow_I n; R \qquad be, S \Rightarrow_B b; R \qquad C, S \Rightarrow_C S'; R, W$$

for $S, S' \in States = (Pvar \to \mathbb{Z})$ and $R, W \subseteq Pvar$. $C, S \Rightarrow_C S'; R, W$ means 'in state S, command C executes to state S' whilst reading the set of variables R and writing the set of variables W'. Similarly, if $e \in Iexp \cup Bexp$ then $e, S \Rightarrow v; R$ means 'in state S, e reads variables R in evaluating to v'.

[5 marks]

(b) For $be \in Bexp$, $ie \in Iexp$, $C \in Com$ use induction on the structure of phrases to give simple definitions of sets

$$\mathcal{R}(ie), \mathcal{R}(be), \mathcal{PR}(C), \mathcal{PW}(C) \ \subseteq \ Pvar$$

where $\mathcal{R}(e)$ is the set of variables accessed by e, $\mathcal{PR}(C)$ is a set of variables possibly read during the execution of C and $\mathcal{PW}(C)$ is a set of variables possibly written to during the execution of C. Give an example to show that it is *not* in general true that

$$C, S \Rightarrow_C S'; R, W \quad \text{implies} \quad W = \mathcal{PW}(C). \qquad \text{[5 marks]}$$

(c) Prove that for any C, S, S', R, W

$$C, S \Rightarrow_C S'; R, W \quad \text{implies} \quad (\forall x \in Pvar.\ x \notin W \Rightarrow S(x) = S'(x))$$

and that for any be, S, S', b, R

$$(\forall x \in \mathcal{R}(be). S(x) = S'(x)) \quad \text{implies} \quad (be, S \Rightarrow_B b; R \iff be, S' \Rightarrow_B b; R)$$

[5 marks]

(d) Prove that for any C_1, C_2, C_3 and be that if $\mathcal{R}(be) \cap \mathcal{PW}(C_1) = \emptyset$ then

$$(C_1; \texttt{if } be \texttt{ then } C_2 \texttt{ else } C_3) \ \approx \ \texttt{if } be \texttt{ then } (C_1; C_2) \texttt{ else } (C_1; C_3)$$

You may assume without proof that if $C, S \Rightarrow_C S'; R, W$ then $W \subseteq \mathcal{PW}(C)$.

[5 marks]

15 Numerical Analysis II

With reference to solution of the differential equation $y' = f(x, y)$, explain the conventional notation x_n, $y(x_n)$, y_n, f_n. [4 marks]

Derive Euler's method

$$y_{n+1} = y_n + hf(x_n, y_n). \tag{1}$$

[3 marks]

Euler's method has *local error*

$$\frac{h^2}{2} y''(\xi).$$

Explain the terms *local error*, *global error*. [2 marks]

The multistep formula

$$y_{n+1} = y_{n-3} + \frac{4h}{3}(2f_n - f_{n-1} + 2f_{n-2}) \tag{2}$$

has *local error*

$$\frac{14}{45} h^5 y^{(5)}(\xi).$$

Outline the technique for deriving multistep formulae such as (2). (Omit algebraic details.) [2 marks]

Suppose Euler's formula is used as a starting procedure for formula (2). How many initial steps of formula (2) need to be evaluated using Euler? [2 marks]

Estimate very roughly the number N of Euler steps needed to approximate f_1. (Assume that $\mid y^{(5)}(x) \mid \simeq 30$, and Euler's method has *global error* h/N.) [5 marks]

What is the most important requirement for a starting procedure? Suggest a more suitable starting procedure than Euler. [2 marks]