

COMPUTER SCIENCE TRIPOS Part II

Tuesday 3 June 1997 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.*

1 Specification and Verification I

The command `REPEAT (E) S` is executed by repeating S exactly n times, where n (assumed greater than zero) is the value of E in the initial state.

Devise a Floyd–Hoare style proof rule for establishing partial correctness specifications of the form $\{P\} \text{REPEAT } (E) S \{Q\}$. [6 marks]

Justify the correctness of your rule informally. [4 marks]

What verification conditions should `REPEAT (E) S` generate? [4 marks]

Prove

$$\{Y \geq 0\} Z := 0; \text{REPEAT } (Y) Z := Z + X \{Z = X \times Y\}$$

[6 marks]

2 VLSI

You are required to implement a serial point-to-point link using application-specific VLSI chips. The speed of the link is to be 1 gigabit/sec and the connection to the two communicating computers is to be 64 bits wide.

Sketch a possible architecture using two chips for each interface. The sketch should indicate the major components within each chip and which current technology is appropriate for each chip. [10 marks]

How many gates and pins do you estimate each chip will contain? Justify your answer on the basis of the design requirements. [10 marks]

3 Comparative Architectures

You are a computer architect working on the design of your company's new instruction set architecture for the 21st Century. Analysis of the latest implementation of the current architecture indicates that removal of logic to perform 8- and 16-bit loads and stores could result in a 20% reduction in processor cycle time. You have been assigned to assess the overall performance implications of removing sub-word memory accesses from the current architecture in order to determine whether they will be omitted from the new architecture.

First, you need to consider what sequence of instructions will be required to emulate sub-word loads and stores in software using standard instructions. Show the instruction sequences that will be required to load and store a *signed byte* value given an arbitrary byte address stored in a register. Be sure to state any assumptions you make. [10 marks]

Here is a summary of some dynamic instruction mix data that have been collected from the company's current processor executing an important integer benchmark:

Instruction	Percentage	Instruction	Percentage
load	20%	add	18%
branch	16%	compare	13%
store	9%	or	8%
shift	7%	other	9%

Of the loads and stores, 94% are 32-bit, 0% are 16-bit, and 6% are 8-bit. Assuming that your instruction sequences are used to replace all the 8-bit memory accesses, estimate the overall performance of the new implementation relative to the one with hardware support for such accesses. [4 marks]

In practice, shorter instruction sequences can be used to replace most sub-word accesses on processors without such hardware support. What extra instruction set features or compiler optimisations might be used to reduce the overhead of sub-word memory accesses? [6 marks]

4 ECAD Topics

Describe the principal data structures and algorithms used in the construction of an event-based digital logic simulator. [10 marks]

Illustrate your answer with a description of how the simulator would model an RS flip-flop made from two cross-coupled NOR-gates with the following inputs:

Time 10	R & S both zero
15	R becomes one
20	S becomes one
25	R becomes zero
30	S becomes zero

The delay of the Q gate should be 6 units and that of the \bar{Q} gate 3. [10 marks]

5 Denotational Semantics

Suppose that for each domain D we wish to have a function $Y_D : (D \rightarrow D) \rightarrow D$ (where as usual, $D \rightarrow D$ denotes the domain of continuous functions from D to itself). Consider the following two properties:

- (a) For all continuous functions $f : D \rightarrow D$, $Y_D(f) = f(Y_D(f))$.
- (b) For all domains E and for all continuous functions $f : D \rightarrow D$, $g : E \rightarrow E$ and $h : D \rightarrow E$, with h strict, if $h \circ f = g \circ h$ then $Y_E(g) = h(Y_D(f))$.

Prove that there is such a family of functions Y_D satisfying (a) and (b). Any facts that you use about the existence of fixed points must be proved. [10 marks]

Let $\Omega = \{d_n \mid n \geq 0\} \cup \{d_\omega\}$ be the domain consisting of a countably infinite chain together with its least upper bound: thus the non-trivial instances of the partial order relation on Ω are $d_0 \sqsubseteq d_1 \sqsubseteq d_2 \sqsubseteq \dots \sqsubseteq d_\omega$. Let $s : \Omega \rightarrow \Omega$ be the continuous function that sends each d_n to d_{n+1} and sends d_ω to itself. For each continuous function $g : E \rightarrow E$ on a domain E , show that there is a strict continuous function $h : \Omega \rightarrow E$ satisfying $h \circ s = g \circ h$. Hence, or otherwise, deduce that there is only one family of functions Y_D satisfying properties (a) and (b). [10 marks]

6 Artificial Intelligence

What kinds of knowledge would an automated travel agent need to represent? [5 marks]

Using the automated travel agent as an example, define *five* problems connected with knowledge representation. [10 marks]

Would an expert system be a suitable technique for implementing an automated travel agent? [5 marks]

7 Neural Computing

Illustrate how stochasticity can be used in artificial neural networks to solve, at least in an asymptotic sense, problems that would otherwise be intractable. Name at least *two* such stochastic engines, describe the rôle of stochasticity in each, and identify the kinds of problems that such artificial neural devices seem able to solve. [10 marks]

Illustrate the evidence for stochasticity in natural nervous systems, and comment on the rôle that it might play in neurobiological function. What is the case supporting John von Neumann's deathbed prediction that stochasticity might be a computational engine for the nervous system, rather than just random noise? Describe at least one experiment involving neural tissue in support of this theory. [10 marks]

8 Database Topics

Describe the basic architecture of the ODMG-93 standard for Object Database Management. [5 marks]

In what way do these proposals allow database management to be integrated with Object-Oriented Distributed Programming? [3 marks]

Explain the properties of collection types, taking care to distinguish between literals and objects. [6 marks]

Describe how an SQL-compatible query language may be defined within the ODMG standard. What provision can be made for insertion, update and deletion? [6 marks]

9 Security

Describe the following modes of operation of a block cipher: electronic codebook, cipher block chaining, output feedback, cipher feedback, message authentication code and hash function. [12 marks]

With which of these modes would it usually be unwise to use the Data Encryption Standard algorithm? [3 marks]

How would you choose a mode of operation to protect the confidentiality of data traffic on a radio link with known rates of (*a*) bit errors and (*b*) burst errors causing loss of synchronisation? [5 marks]

10 Types

Let x range over a set of identifiers, i range over integer constants, r range over real constants and α range over a set of type variables. Now suppose we have a set of types, σ , given by

$$\sigma ::= \alpha \mid \text{int} \mid \text{real} \mid \sigma_1 \times \sigma_2 \mid \sigma_1 \rightarrow \sigma_2$$

and a language of terms, M , given by

$$M ::= x \mid i \mid r \mid \lambda x.M_1 \mid M_1 M_2 \mid (M_1, M_2).$$

Give ML-like type inference rules for formulae of the form $\Gamma \vdash M : \sigma$, explaining the form and rôle of Γ . [4 marks]

Explain the notion of *principal type* and state whether your set of rules has such a property. [4 marks]

Show from your rules that it is impossible to find a Γ and σ which enable inference of either $\Gamma \vdash \lambda f.(f(1), f(2.7)) : \sigma$ [2 marks]
or $\Gamma \vdash \lambda x.xx : \sigma$. [2 marks]

Now suppose we wish to do better than the usual ML treatment of overloading for operators like “+”. So add to the language of types a conjunction connective

$$\sigma ::= \sigma_1 \wedge \sigma_2$$

where $M : \sigma_1 \wedge \sigma_2$ means informally that M has both types σ_1 and σ_2 and so can be used at either type. Add corresponding inference rules:

$$\frac{\Gamma \vdash M : \sigma \quad \Gamma \vdash M : \sigma'}{\Gamma \vdash M : \sigma \wedge \sigma'}$$

$$\frac{\Gamma \vdash M : \sigma \wedge \sigma'}{\Gamma \vdash M : \sigma} \qquad \frac{\Gamma \vdash M : \sigma \wedge \sigma'}{\Gamma \vdash M : \sigma'}$$

Now show that a suitable σ_1 and σ_2 can be found such that

$$[\text{neg} : (\text{int} \rightarrow \text{int}) \wedge (\text{real} \rightarrow \text{real})] \vdash \lambda f.(f(\text{neg}(1)), f(\text{neg}(2.7))) : \sigma_1$$

and

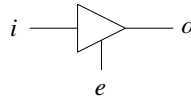
$$[] \vdash \lambda x.xx : \sigma_2$$

hold. [8 marks]

[Hint: for the latter, you might give x a type $\sigma \wedge \sigma'$ where σ is a function type which accepts σ' as an argument.]

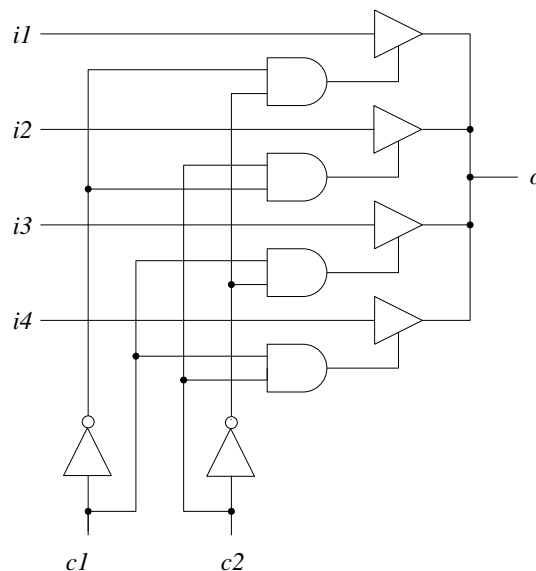
11 Specification and Verification II

A tri-state buffer connects its input i to its output o when the enable line e is driven with Hi. When e is Lo the output o is in a high impedance state Z .



Write down and explain a logical predicate modelling such a buffer. [4 marks]

The circuit shown below is a four-way multiplexer designed with tri-state logic. Assume that the control inputs $c1$, $c2$ and the data inputs $i1$, $i2$, $i3$, $i4$ are either Hi or Lo.



Write down a specification of the multiplexer. [4 marks]

Write down predicates modelling an inverter and an AND-gate. [2 marks]

Explain how the multiple driving of the output o by the four buffers can be modelled. [4 marks]

Explain how to compose together the predicates modelling the components of the multiplexer into a predicate representing the complete circuit and write down the result. [6 marks]

12 Communicating Automata and Pi Calculus

Define the concepts of *strong simulation* and *strong bisimulation* between processes. [4 marks]

Here are three vending machines, selling tea for 20p and/or coffee for 40p:

$$\begin{aligned} T &\stackrel{\text{def}}{=} 20p.tea.T \\ C &\stackrel{\text{def}}{=} 20p.20p.coffee.C \\ CT &\stackrel{\text{def}}{=} 20p.(tea.CT + 20p.coffee.CT) \end{aligned}$$

“20p” is the user’s action of inserting a coin; “tea”, “coffee” are the user’s actions of selecting a drink. (For simplicity, the delivery of a drink is not represented.)

One might hope that CT can be made to behave like C or T respectively by blocking off (i.e. by restricting by ν) the appropriate select button. But prove, by considering bisimulations or otherwise, that one of the following assertions is true, the other false:

$$\begin{aligned} (\nu tea) CT &\sim C \\ (\nu coffee) CT &\sim T \end{aligned} \quad [8 \text{ marks}]$$

In the case which is false, is there a simulation in one direction? Give a precise reason for your answer. [3 marks]

Now redesign T , C and CT , in a way which still has reasonable behaviour for a machine selling tea and/or coffee, but so that *both* of the above assertions hold, and justify the claim. [5 marks]

13 Natural Language Processing

The following dialogue between a hypothetical User, U, and a database access system, S, illustrates some of the problems inherent in natural language processing by machine.

U: Is every type of shock absorber stocked in one warehouse in any city?
 S: Yes, in the South Coventry warehouse.
 U: Which types are held in Birmingham?
 S: sa55 and sa65 shock absorbers for Metros and Montegos.

- (a) Describe the problems illustrated by this extract. [6 marks]
- (b) Give a sketch of an architecture of a natural language interface to a database which could support such exchanges. [6 marks]
- (c) Say how, and to what extent, the problems described could be solved within such an architecture. [8 marks]

14 Additional Topics

Explain how to initiate and operate end-to-end data streams in ATM networks, including both cell-level and call-level issues. [7 marks]

List *two* problems with ATM in the wireless/mobile environment. [6 marks]

List *three* factors which affect the performance of TCP over ATM, and explain how they affect throughput. [7 marks]

15 Additional Topics

Discuss the motives which lead people to design programming languages. Illustrate your discussion with languages which were apparently designed from differing motives. [10 marks]

Discuss the criteria which may be used to measure the success of a programming language. How does Standard ML fare, according to these criteria? How does its type system contribute to its success or failure? [10 marks]