

## COMPUTER SCIENCE TRIPOS Part II

---

Thursday 7 June 2001 1.30 to 4.30

---

Paper 9

*Answer **five** questions.*

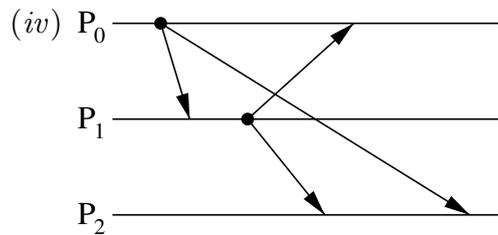
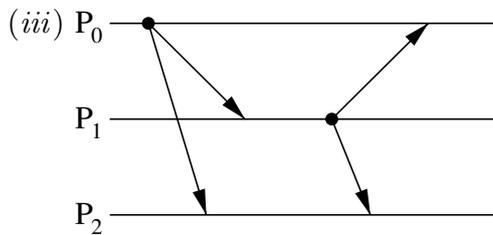
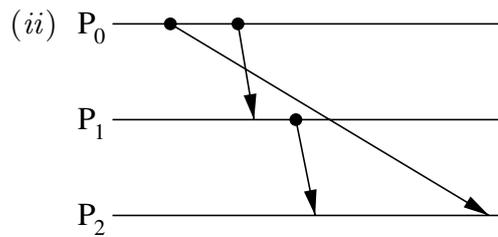
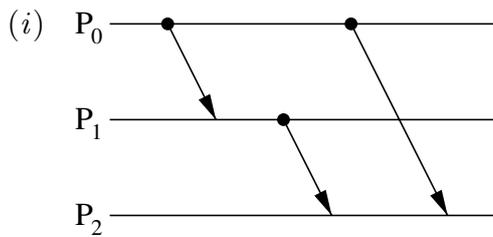
*Submit the answers in five **separate** bundles, each with its own cover sheet. On each cover sheet, write the numbers of **all** attempted questions, and circle the number of the question attached.*

*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 Distributed Systems

- (a) Explain the problem of clock drift in distributed systems. [2 marks]
- (b) What sources of conventional earth time might be used by computer systems? How would you estimate bounds on the accuracy of time received from such a source? [4 marks]
- (c) What constraint does distributed inter-process communication (IPC) impose on the clock values of the communicating parties? [1 mark]
- (d) Outline one clock synchronisation protocol that satisfies this constraint. [4 marks]
- (e) For each of the cases of IPC illustrated below, give the vector clock values that message receiving and delivery modules could maintain for each process.

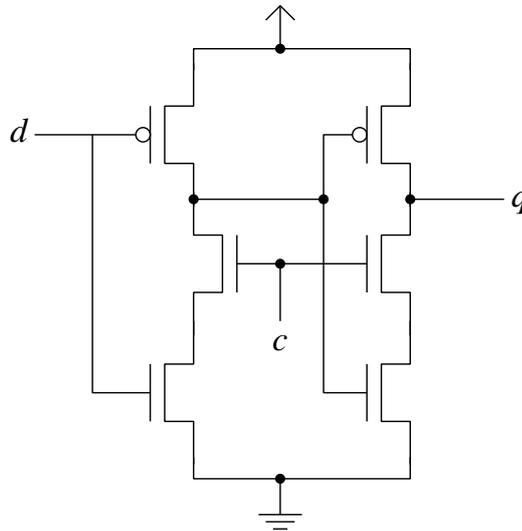


[6 marks]

- (f) Define “causal order” of message delivery. In which, if any, of (i) to (iv) above is causal order violated at the message receiving module? [3 marks]

## 2 VLSI Design

- (a) Give transistor level designs for 2-input NAND gates using static and dynamic CMOS, explaining how the latter is controlled by a clock. [8 marks]
- (b) Comment on the relative merits of static and dynamic logic for evaluating more complicated combinatorial functions, including transistor count, wiring complexity, speed and implications for cascaded logic. [4 marks]
- (c) Consider the following circuit for a dynamic latch:



Specify the input and output signals, and explain its operation carefully.

[8 marks]

### 3 Digital Communication II

Write notes about *four* of the following types of Access Network. You should discuss any special challenges posed by each transmission medium, describe the current technology, and comment on any commercial factors pertinent to its deployment.

- (a) Analogue telephone modems.
- (b) Cable modems.
- (c) Digital Subscriber Line (xDSL).
- (d) Fixed wireless.
- (e) Optical fibre to the home.
- (f) Satellite.

[5 marks each]

#### 4 Advanced Graphics and HCI

(a) Which usability evaluation techniques are most appropriate to the following situations? For *each*, say why this is so.

(i) An expert user performing a familiar task.

(ii) An expert user designing a novel solution.

(iii) A new user using a system for the first time.

[6 marks]

(b) Sketch *three* screens of a prototype for an online newspaper that can be customised to users' interests.

(i) How could a development team test the usability of this prototype cheaply and quickly? Describe some basic procedures that could be applied during testing.

(ii) Contrast the benefits of doing a Cognitive Walkthrough evaluation to those of the procedure you have just described, in the context of this system.

(iii) Give an example of one stage of the Cognitive Walkthrough process, as it would apply to one of the screens you have sketched.

[14 marks]

## 5 Business Studies

- (a) What is meant by the terms *supply curve* and *demand curve*, and what is significant about the point where they cross? [5 marks]
- (b) What is meant by the terms *cost curve* and *break-even point*? Draw a diagram to show the relation between them and the demand curve. How would you attempt to establish a demand curve in practice? [5 marks]
- (c) The cost and demand schedules for a particular product are given in the following table. What price should the manufacturer set? [5 marks]

Volume, $k$	Unit Cost	Unit Price
1	11.8	19
2	11.6	18
3	11.4	17
4	11.2	16
5	11.0	15
6	10.8	14
7	10.6	13
8	10.4	12
9	10.2	11
10	10.0	10

- (d) Discuss the economic impact of the advent of the Internet. [5 marks]

## 6 Types

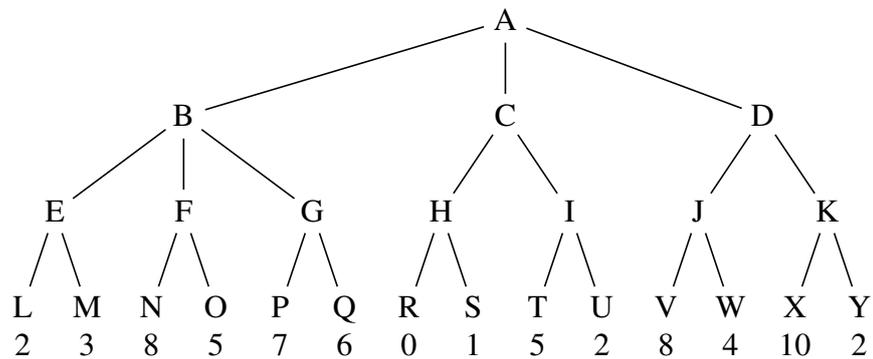
- (a) Define the typing relation of the polymorphic lambda calculus (PLC). [5 marks]
- (b) For *each* of the following PLC typing judgements, state whether or not there is a type  $\tau$  that makes the judgement provable. Justify your answer in each case.
- (i)  $\vdash \lambda x : \forall \alpha (\alpha) (\Lambda \beta (x \beta)) : \tau$  [3 marks]
- (ii)  $\vdash \Lambda \alpha (\lambda x : \alpha (\Lambda \beta (x \beta))) : \tau$  [3 marks]
- (iii)  $\vdash \lambda x : \tau (\Lambda \alpha (x (\alpha \rightarrow \alpha) (x \alpha))) : \tau \rightarrow \forall \beta (\beta)$  [3 marks]
- (iv)  $\vdash \lambda x : \tau (\Lambda \alpha (x (\alpha \rightarrow \alpha) (x \alpha))) : \tau \rightarrow \forall \alpha (\alpha \rightarrow \alpha)$  [3 marks]
- (v)  $\vdash \Lambda \alpha (\lambda x : \tau (x (\alpha \rightarrow \alpha) (x \alpha))) : \forall \alpha (\alpha \rightarrow \alpha)$  [3 marks]

## 7 Optimising Compilers

- (a) Sketch an algorithm for *instruction scheduling*. Be careful to specify the nature of your initial and final data-structures; also give the worst-case running time of the algorithm and a program which causes this worst-case running time to be achieved. [10 marks]
- (b) Explain the conflicts between register allocation, common sub-expression elimination and instruction scheduling on a machine where register usage should be minimised (for example to reduce procedure entry save/restore code). [5 marks]
- (c) Describe Static Single Assignment (SSA) form and explain how it can help to produce better results in both compilation and decompilation. [5 marks]

## 8 Artificial Intelligence

- (a) Why does search in game-playing programs always proceed forward from the current position rather than backward from the goal? [5 marks]
- (b) Most game-playing programs do not save search results from one move to the next. Instead, they usually start completely afresh whenever it is the machine's turn to move. Why? [5 marks]
- (c) Consider the following game tree, and assume that the first player is the maximising player:



- (i) Which move should the first player choose? [3 marks]
- (ii) Assuming that nodes are searched left-to-right using the alpha-beta algorithm, which nodes would not be examined? [7 marks]

## 9 Neural Computing

- (a) (i) Many classes of artificial neural networks learn from data by forming a lower dimensional parametric representation, or mapping, that resembles the process of curve-fitting. Explain this idea in reference to least-squares error minimisation or statistical regression. [4 marks]
- (ii) Explain why increasing the complexity of a model may cause a neural network's error in the training phase to become smaller and smaller, but its generalisation in the validation phase to become worse and worse. How would you expect the optimal complexity of a neural network model to depend on the amount of data? [6 marks]
- (b) Answer *each* of the following short questions:
- (i) What is the approximate capacitance of nerve cell membrane, in microFarads per  $\text{cm}^2$ , and what functional parameters of neural activity are determined by this? [2 marks]
- (ii) Approximately what range of voltages does a nerve cell membrane move through during the course of generating a neural impulse, and what determines this range? [2 marks]
- (iii) What is the rôle of positive feedback in nerve impulse generation? [2 marks]
- (iv) From which organ does the retina develop embryologically, and to what cells elsewhere in the body are the retinal photoreceptors most closely related? [2 marks]
- (v) What causes the refractory deadtime of about 1 msec after each nerve impulse, and what is its consequence? [2 marks]

## 10 Information Theory and Coding

- (a) (i) Construct an efficient, uniquely decodable binary code, having the prefix property and having the shortest possible average code length per symbol, for an alphabet whose five letters appear with these probabilities:

<i>Letter</i>	<i>Probability</i>
A	1/2
B	1/4
C	1/8
D	1/16
E	1/16

[4 marks]

- (ii) How do you know that your code has the shortest possible average code length per symbol? [2 marks]
- (b) (i) For a string of data of length  $N$  bits, what is the upper bound for its Minimal Description Length, and why? [2 marks]
- (ii) Comment on how, or whether, you can know that you have truly determined the Minimal Description Length for a set of data. [2 marks]
- (c) Suppose you have sampled a strictly bandlimited signal at regular intervals more frequent than the Nyquist rate; or suppose you have identified all of the zero-crossings of a bandpass signal whose total bandwidth is less than one octave. In either of these situations, provide some intuition for why you now also have knowledge about exactly what the signal must be doing at all points between these observed points. [3 marks]
- (d) (i) Explain how autocorrelation can remove noise from a signal that is buried in noise, producing a clean version of the signal. For what kinds of signals, and for what kinds of noise, will this work best, and why? What class of signals will be completely unaffected by this operation except that the added noise has been removed? Begin your answer by writing down the autocorrelation integral that defines the autocorrelation of a signal  $f(x)$ . [5 marks]
- (ii) Some sources of noise are additive (the noise is just superimposed onto the signal), but other sources of noise are multiplicative in their effect on the signal. For which type would the autocorrelation clean-up strategy be more effective, and why? [2 marks]

## 11 Numerical Analysis II

(a) Taylor's theorem states that if  $x \in [a, b]$  and  $f \in C^{N+1}[a, b]$

$$f(x) = T_N(a) + \frac{1}{N!} \int_a^x f^{(N+1)}(t)(x-t)^N dt$$

where

$$T_N(a) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \dots + \frac{(x-a)^N}{N!}f^{(N)}(a).$$

Prove Taylor's theorem.

[6 marks]

(b) Peano's theorem states that if a quadrature rule integrates polynomials of degree  $N$  exactly over an interval  $[a, b]$  then the error in integrating  $f \in C^{N+1}[a, b]$  can be expressed as

$$E(f) = \int_a^b f^{(N+1)}(t)K(t) dt$$

where

$$K(t) = \frac{1}{N!} E_x[(x-t)_+^N].$$

Explain the notation  $E(f)$ ,  $E_x$  and  $(x-t)_+^N$ .

[4 marks]

(c) Use Taylor's theorem to prove Peano's theorem.

[8 marks]

(d) Under what additional condition may the simplified formula

$$E(f) = \frac{f^{(N+1)}(\xi)}{(N+1)!} E(x^{N+1})$$

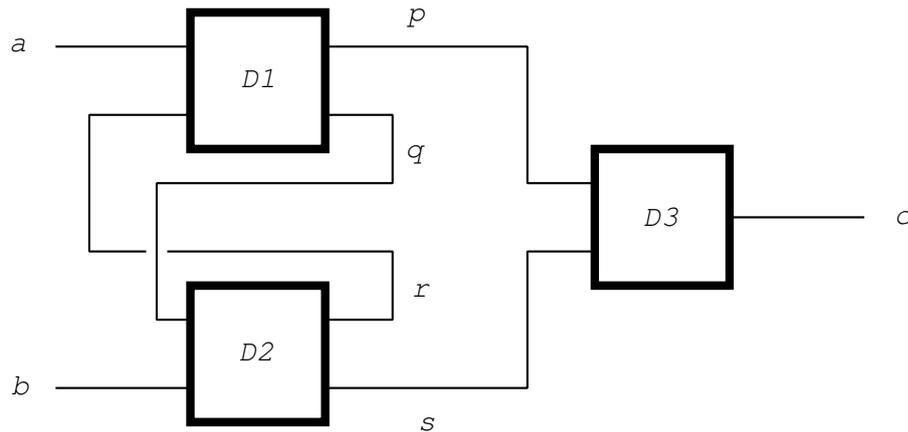
be applied?

[2 marks]

## 12 Specification and Verification II

Suppose definitions of  $D1(a, r, p, q)$ ,  $D2(q, b, r, s)$  and  $D3(p, s, c)$  are given.

- (a) Using  $\exists$  and  $\wedge$  write down a definition of a predicate  $D$  such that  $D(a, b, c)$  defines the relation between  $a$ ,  $b$  and  $c$  when  $D1$ ,  $D2$  and  $D3$  are connected as in the following diagram. [4 marks]



Suppose now that  $D1$ ,  $D2$  and  $D3$  are defined in terms of functions  $f_1, f_2, f_3, f_4$  and  $f_5$  by

$$\begin{aligned} D1(a, r, p, q) &= (p = f_1 a r) \wedge (q = f_2 a) \\ D2(q, b, r, s) &= (r = f_3 q b) \wedge (s = f_4 q b) \\ D3(p, s, c) &= (c = f_5 p s) \end{aligned}$$

- (b) Write down an equation expressing  $c$  in terms of  $a$  and  $b$ . [4 marks]
- (c) Show the logical steps needed to derive the equation from the definition of  $D$ . [12 marks]

## 13 Computer Vision

Give *three* examples of problems in computer vision which are formally ill-posed. In each case explain how one or more of Hadamard's criteria for well-posed problems has failed to be satisfied. Illustrate how addition of ancillary constraints or assumptions, even metaphysical assumptions about how the world behaves, enable one to convert the ill-posed problem into a well-posed problem. Finally, discuss how the use of Bayesian priors can perform this function.

[20 marks]

## 14 Natural Language Processing

“Kim ran up a hill by the steepest path. Sandy had egged her on.”

This narrative exemplifies some problems for natural language processing systems.

- (a) Identify these problems. [6 marks]
- (b) Describe the stages of analysis that would be required to interpret the narrative. [6 marks]
- (c) Evaluate how successful each stage is likely to be. [8 marks]

## 15 Topics in Concurrency

This question assumes familiarity with the process language **SPL** and its event-based semantics. In the following **SPL** process, *Auth*, agents can behave as initiator or responder in parallel with an attacker *Spy*. Letting *A* and *B* range over agent names, define

$$Init(A, B) \equiv out\ new\ x\ \{x, A\}_{Pub(B)}.in\ \{x\}_{Pub(A)}.nil$$

$$Resp(B) \equiv in\ \{x, X\}_{Pub(B)}.out\ \{x\}_{Pub(X)}.nil$$

$$Auth \equiv \parallel_{i \in \{init, resp, spy\}} P_i \quad \text{where}$$

$$P_{init} \equiv \parallel_{A, B} Init(A, B), \quad P_{resp} \equiv \parallel_A !Resp(A), \quad \text{and} \quad P_{spy} \equiv !Spy$$

- (a) Explain briefly and informally the behaviour of *Init(A, B)* and *Resp(B)*, for agent names *A* and *B*. Describe diagrammatically the reachable events of *Init(A, B)* and *Resp(B)*, taking care to specify the pre- and postconditions, and actions of the events. [5 marks]
- (b) Write down an **SPL** process for the attacker *Spy*; the process should be able to compose, decompose, encrypt under public keys, and decrypt with leaked private keys. Draw the reachable events of *Spy*. [5 marks]

Assume a sequence of event-transitions

$$\langle Auth, s_0, t_0 \rangle \xrightarrow{e_1} \cdots \langle p_{r-1}, s_{r-1}, t_{r-1} \rangle \xrightarrow{e_r} \langle p_r, s_r, t_r \rangle \cdots$$

from the configuration  $\langle Auth, s_0, t_0 \rangle$ , of which it is assumed that the names in *Auth* and the output messages  $t_0$  are included in the name-set  $s_0$ . Suppose that the event  $e_r$  is the input of a message  $\{m\}_{Pub(A)}$  by agent *A* as initiator. Define a property of subsets of messages  $t$  by

$$Q(t) \text{ iff } \forall M \in t. m \sqsubset M \Rightarrow \{m, A\}_{Pub(B)} \sqsubset M,$$

where, for instance,  $m \sqsubset M$  means  $m$  is a submessage of  $M$ .

- (c) Explain briefly why  $Q(t_0)$  is true and  $Q(t_{r-1})$  is false. [6 marks]
- (d) Describe, without proof, the possible form(s) of the earliest event  $e_i$  for which  $Q(t_{i-1})$  is true while  $Q(t_i)$  is false. [4 marks]

**END OF PAPER**