COMPUTER SCIENCE TRIPOS Part II

Wednesday 3 June 1998 1.30 to 4.30

Paper 8

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

Give a detailed proof using the method of verification conditions that:

$\vdash \{\texttt{N} = n \ \wedge \ n \geqslant 0\}$

```
FACT := 1;
WHILE N>1 DO
  X := FACT; Y := N;
  WHILE Y>1 DO
    FACT := FACT+X; Y := Y-1
  N := N-1
```

$\{\texttt{FACT} = n!\}$

You should write out an annotated specification and carefully list the verification conditions. [20 marks]

2 Digital Communication II

"A TCP implementation tries to model the state of the receiver and of the channel."

(a) Describe in detail which aspects of the receiver state and the channel are being modelled. [5 marks]

(b) To what extent can the model which TCP builds up be inaccurate? What are the consequences of the inaccuracy? [5 marks]

(c) "TCP works well for long-lived connections but not so well for short-lived connections." Discuss. [5 marks]

(d) "Load balancing of links in the internet is a mixed blessing." Discuss. [5 marks]

3 Computer System Modelling

Define the term *Markov Chain*. Why is the Markov property useful in modelling queueing systems? [5 marks]

Consider a birth–death queueing system with the following birth and death coefficients in which the state index represents the number of customers in the system:

$$\begin{aligned} \lambda_k &= (k+2)\lambda \qquad k = 0, 1, 2 \ldots \\ \mu_k &= k\mu \qquad\qquad k = 1, 2 \ldots \end{aligned}$$

All other coefficients are zero. Solve for p_k, the set of equilibrium probabilities for all states k, for $k = 0, 1, 2 \ldots$ State how you would find the average number of customers in the system. [15 marks]

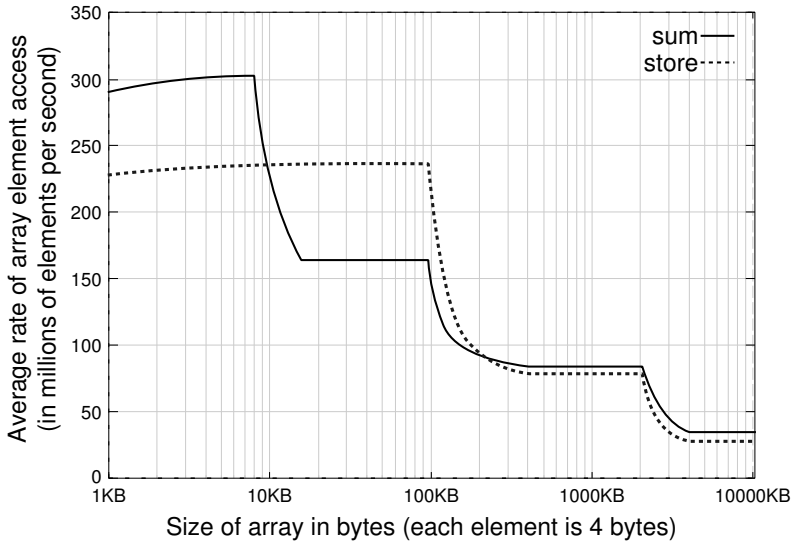
4 Comparative Architectures

Modern workstations typically have memory systems that incorporate two or three levels of caching. Explain why they are designed like this. [4 marks]

In order to investigate the performance of a system's memory hierarchy, a user writes a simple test program. The program, known as `sum`, accepts N as an input parameter, and allocates an array of N 32-bit integer elements in physically contiguous memory. `sum` contains an inner loop that scans sequentially over the array and computes the sum of all the elements. The program measures the total time taken to execute several thousand iterations of the scan, and uses this to compute the average rate at which the computation proceeded.

Write sample assembly code for a performance-optimised implementation of `sum`'s inner loop for a super-scalar RISC processor. You may assume that the array always contains a multiple of eight elements, and you are encouraged to demonstrate your knowledge of techniques such as loop unrolling and instruction scheduling. Indicate how your loop might execute on a processor capable of issuing two integer operations per cycle. [8 marks]

In addition to `sum`, the user writes a similar program called `store`. Instead of totalling the array, `store` simply writes to each element, setting its contents to zero. The user invokes the programs over different sizes of array in order to generate the following graph. The x axis indicates the size of the array (in bytes) that the program was operated over, while the y axis shows the average rate at which array elements were processed (in millions per second).



Ignoring startup effects, describe the behaviour of the memory system, and hence account for the graph. What can you deduce about the workstation's memory hierarchy? [8 marks]

5 Business Studies

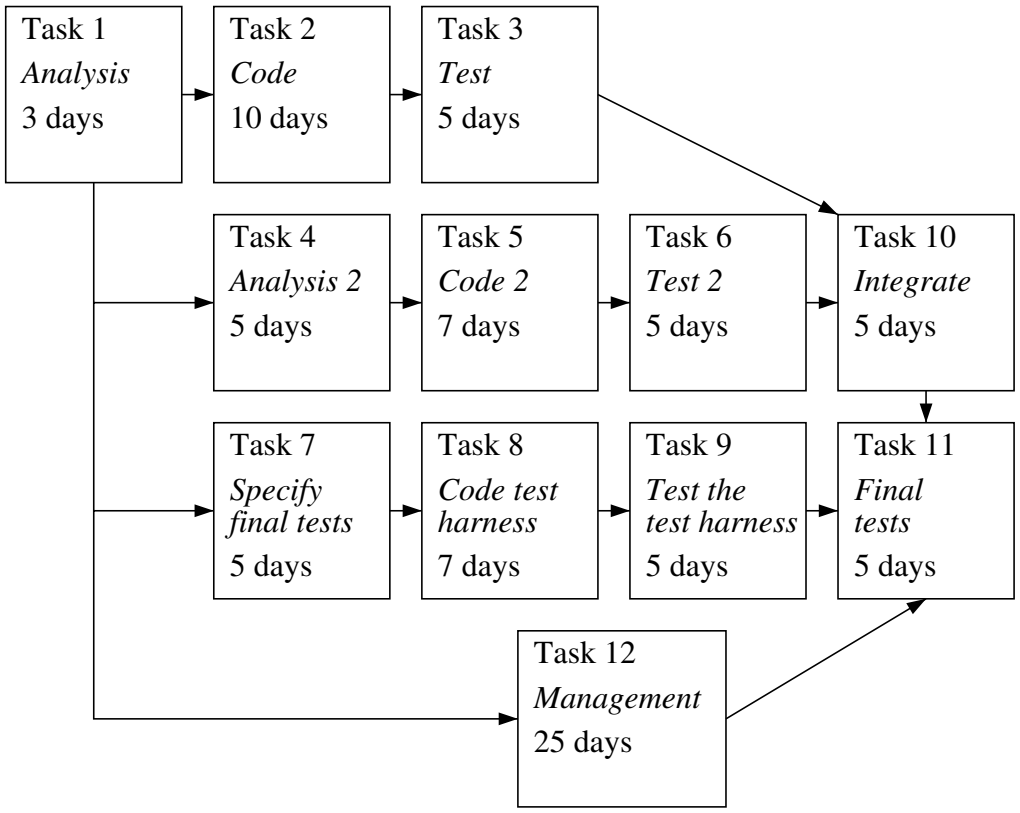
The figure shows a section of a PERT diagram for a small software project, together with the number of programmer-days each task is estimated to take.

(*a*) What is the *critical path* in a PERT diagram and why is it important? What is the critical path in the PERT diagram shown below? [5 marks]

(*b*) One analyst and two programmers are assigned to the project, in addition to the manager. How long will the project take? [5 marks]

(*c*) Derive the equivalent GANTT chart for the project. [5 marks]

(*d*) Task 8 slips by 2 days. How does this affect the project? What actions, if any are required, can be taken to alleviate the slippage? [5 marks]



6 Advanced Algorithms

Suppose you are presented with a (large) integer N and are asked to find its complete factorisation. You are not told anything at all in advance about how many factors it will have, but you are instructed to use the Pollard Rho method as a probabilistic algorithm as the core of your code. Explain

(a) The overall structure of the code you would write, where it calls Pollard Rho and any other sub-algorithms you will use, and what their purpose is. [4 marks]

(b) What steps are taken in the Rho method and (informally) why it might be hoped that it will do what it is expected to. [6 marks]

(c) The extents and manners in which parts of your code rely on random numbers and the consequences of these turning out to be either especially fortunate or especially awkward. [5 marks]

(d) A coarse estimate of the total run-time for the factoriser in circumstances when the input number has exactly two large prime factors, and an equally crude estimate of the size that N would need to be before the factorisation process took a whole day of CPU time on a modern desktop workstation. You may suppose that around 10^{13} basic operations are available in that amount of time. [5 marks]

7 Optimising Compilers

(a) Define the notion of an expression being available at a node in a flowgraph in terms of possible execution flows of control; explain carefully the form which available expressions might take in your framework. [4 marks]

(b) Demonstrate that calculating exactly which expressions are available at a given node is uncomputable (you may assume that it is uncomputable to determine whether two given boolean expressions involving arithmetic always yield the same result value). [4 marks]

(c) Give an algorithm to calculate available expressions and state carefully how the algorithmic result is related to the set of expressions which are available according to your definition in part (a). [4 marks]

(d) Justify any discrepancy in (c) by reference to safety of the approximation with respect to the usual use of available expressions in optimisation. [4 marks]

(e) On a machine with four registers available for register allocation (by colouring), give a program for which common sub-expression elimination (CSE) results in worse code being generated than if CSE had not been performed, noting any assumptions of timing factors for the target machine which justify the code being worse. [4 marks]

8 Artificial Intelligence

Give short definitions of any *ten* of the following terms:

(a) Circumscription

(b) Situation Calculus

(c) Closed World Assumption

(d) Expert System

(e) Automatic Discovery of Heuristics

(f) Inter-Wave Search

(g) Abduction

(h) Unique Names Assumption

(i) Microworld

(j) Maximum Expected Utility

(k) Intelligent Agent Architecture

(l) Fuzzy Control

(m) Constraint

(n) Problem Space (or Search Space)

(o) Non-Monotonic Reasoning

[2 marks each]

9 Database Topics

What are the aims of the Object Database Standard ODMG 2.0? [3 marks]

Describe the ways in which the Standard contributes to Object-Orientated Distributed Programming, making particular reference to distribution and heterogeneity. [5 marks]

What benefits does conformance to the Standard bring to the application developer? It may be helpful to consider

(a) schema definition and maintenance

(b) transaction management

(c) object interchange format

(d) interoperation with relational data storage [12 marks]

10 Information Retrieval

You work for a company that takes news stories from all over the world and provides reports on specified topics to customers, for example on political developments in the new Republic of Rumbaza during 1997. Your company's staff use a retrieval system to extract the material on which they base their reports from the company's very extensive archive of stories. The retrieval system is a 20-year-old makeshift and is to be scrapped. You are asked to design the new retrieval system.

(a) Give a detailed description of the retrieval devices the new system will offer, explaining why they are available and how they will work. What facilities will the user have for modifying his or her search specification in response to system output? [15 marks]

(b) What do you regard as the most difficult problem to be tackled, and why? [5 marks]

[Assume that the stories in the file are texts between 1 and 100 sentences long, with date and source headers in a standardised form.]

11 Information Theory and Coding

Consider a binary symmetric communication channel, having source alphabet $X = \{0, 1\}$ with probabilities $\{0.5, 0.5\}$. Its output alphabet is $Y = \{0, 1\}$ and its channel matrix is

$$\begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix}$$

where ϵ is the probability of transmission error.

(a) What is the entropy of the source, $H(X)$? [1 mark]

(b) What is the probability distribution of the outputs, $p(Y)$, and the entropy of this output distribution, $H(Y)$? [3 marks]

(c) What is the joint probability distribution for the source and the output, $p(X, Y)$, and what is the joint entropy, $H(X, Y)$? [4 marks]

(d) What is the mutual information of this channel, $I(X; Y)$? [2 marks]

(e) How many values are there for ϵ for which the mutual information of this channel is maximal? What are those values, and what then is the capacity of such a channel in bits? [3 marks]

(f) For what value of ϵ is the capacity of this channel minimal? What is the channel capacity in that case? [2 marks]

The Fourier transform (whether continuous or discrete) is defined in the general case for complex-valued data, which gets mapped into a set of complex-valued Fourier coefficients. However, we are often concerned with purely real-valued data, such as sound waves or images, whose Fourier transforms we would like to compute. What simplification occurs in the Fourier domain as a consequence of having real-valued, rather than complex-valued, data? [5 marks]

12 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 the addition of ancillary constraints or assumptions, even metaphysical assumptions about how the world behaves, enables 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]

13 Specification and Verification II

A simple device CP is informally specified as follows:

> CP stores values of type σ. It has two inputs: `inp` carries values of type σ and `func` carries 2-bit words; it has an output line `outp` carrying the value stored. CP is a state machine whose next state depends on the value input at `func`:
>
> - if `func`=0 the value stored is unchanged
> - if `func`=1 the value on `inp` replaces the value stored
> - if `func`=2 the value stored is transformed by a function $f_2 : \sigma \to \sigma$
> - if `func`=3 the value stored is transformed by a function $f_3 : \sigma \to \sigma$

Define a predicate CP that formalises this specification in higher order logic. [4 marks]

Write down logical models of the following components:

- a combinational multiplexer that routes one of four σ-valued inputs to a single output, depending on the value of a 2-bit control input; [3 marks]
- a unit-delay register that holds values of type σ. [3 marks]

Draw a schematic diagram of an implementation of CP built from these components. [3 marks]

Write down a formula that expresses the correctness of your implementation. [4 marks]

Discuss briefly how you would go about proving your correctness formula. You need not give a detailed proof, but you should aim to convince the reader that given time you could produce one. [3 marks]

14 Numerical Analysis II

Let n_+ be the number of positive real roots of a polynomial $p_n(x)$. Let c be the number of changes of sign when the coefficients are taken in order. State *Descartes' rule of signs.* [2 marks]

If

$$p_3(x) = x^3 - 8x^2 + 11x + 20$$

what does this rule say about the polynomials $p_3(x)$, $p_3(-x)$? [2 marks]

Newton's method for solution of a system of n non-linear equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ can be expressed in the form

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_{k+1}.$$

What is the formula for $\mathbf{h}_{k+1}$? [2 marks]

Outline the *damped Newton method* in n variables. [4 marks]

Apply the damped Newton method in one variable to find one root of the polynomial $p_3(x)$, using $x_0 = 1$ as the starting value. Hence factorise $p_3(x)$ and draw a rough sketch of the polynomial, showing the first Newton step as a tangent. [10 marks]

15 Communicating Automata and Pi Calculus

A process is *deterministic* if, for any state P and any action a, there is at most one transition $P \xrightarrow{a} P'$. If $s = a_1 \cdots a_n$ is a sequence of actions, then $P \xrightarrow{s} P'$ means $P \xrightarrow{a_1} \cdots \xrightarrow{a_n} P'$.

By drawing a transition graph, define a deterministic sequential process Q which uses the four actions $\mathsf{in}_1, \mathsf{out}_1, \mathsf{in}_2, \mathsf{out}_2$ such that, for every action sequence s, $Q \xrightarrow{s} Q'$ for some Q' if and only if s satisfies the four conditions

$$(a)\quad 0 \leqslant \#_s\mathsf{in}_1 - \#_s\mathsf{out}_1 \leqslant 1 \qquad (c)\quad 0 \leqslant \#_s\mathsf{in}_1 - \#_s\mathsf{in}_2 \leqslant 1$$
$$(b)\quad 0 \leqslant \#_s\mathsf{in}_2 - \#_s\mathsf{out}_2 \leqslant 1 \qquad (d)\quad 0 \leqslant \#_s\mathsf{out}_1 - \#_s\mathsf{out}_2 \leqslant 1$$

where $\#_s a$ is defined to be the number of times a occurs in s. For example the sequence $s = \mathsf{in}_1\,\mathsf{in}_2\,\mathsf{out}_1$ satisfies the conditions, but $s = \mathsf{in}_1\,\mathsf{in}_2\,\mathsf{out}_2$ violates (d). Also write down the defining equations of Q.
[Hint: Q need have no more than eight states.] [8 marks]

Define *weak bisimulation* and *weak equivalence* ($\approx$), and explain their advantages over the strong versions. [4 marks]

With the help of one or more extra actions $\vec{a}$, define two communicating processes R_1 and R_2 with no more than four states each, such that

$$Q \approx \mathsf{new}\,\vec{a}(R_1 \mid R_2)$$

and verify this weak equivalence. [8 marks]