COMPUTER SCIENCE TRIPOS Part II

Wednesday 2 June 1999 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 Distributed Systems

A distributed database is modelled as distributed persistent objects. Each transaction is submitted to a single node which assumes the role of coordinator, responsible for managing and committing the transaction.

(*a*) Outline how each of the following approaches to concurrency control is implemented for distributed objects:

(*i*) strict two-phase locking [4 marks]

(*ii*) strict timestamp ordering [4 marks]

(*iii*) optimistic concurrency control [4 marks]

(*b*) Explain how atomic commitment of transactions is achieved in a system employing *one* of the above methods of concurrency control. [8 marks]

2 Specification and Verification I

Give an example of P, C and Q such that $\vdash \{P\}\ C\ \{Q\}$ but it is not the case that $\vdash [P]\ C\ [Q]$. Justify your answer. [5 marks]

Explain how to translate the partial correctness specification shown below into higher-order logic.

$$\{\texttt{X} = x\ \wedge\ \texttt{Y} = y\}\ \texttt{TEMP := X;}\ \ \texttt{X := Y;}\ \ \texttt{Y := TEMP}\ \{\texttt{X} = y\ \wedge\ \texttt{Y} = x\}$$

[5 marks]

Write down and justify an example of a correctly annotated specification $\{P\}\ C\ \{Q\}$ such that $\vdash \{P\}\ C\ \{Q\}$ but the verification conditions are not true. Comment on the significance of your example. [5 marks]

Define the meaning of the notation $[P, Q]$. Write down the `WHILE`-law of refinement and justify it with respect to the `WHILE`-rule for Floyd–Hoare logic. [5 marks]

3 Digital Communication II

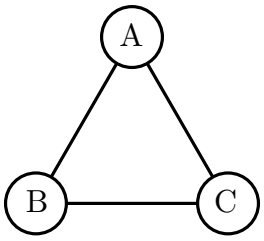
Compare the Integrated Services (IntServ) and Differential Services (DiffServ) Internet communication architectures with respect to

(a) state within the network

(b) the kinds of guarantees that can be made

(c) mechanisms which routers must provide [10 marks]

It is required to provision a network to meet service level agreements (SLAs) which specify traffic requirements. A *single-ended* SLA is one in which the sources but not the destinations of traffic stream requirements are specified, while a *double-ended* SLA specifies both sources and destinations. What are the implications of single-ended and double-ended SLAs on provisioning and the nature of the guarantees that can be made? Use the following as an example.



Case 1: Single-ended SLA

Source	Required rate
A	2
B	3
C	4

Case 2: Double-ended SLA

Source	Destination	Required rate
A	B	1
A	C	1
B	A	2
B	C	1
C	A	2
C	B	2

[10 marks]

4 Comparative Architectures

A new microprocessor has a 64 Kbyte write-back L1 D-cache that is 2-way set-associative (employing a random replacement policy) with 64-byte cache lines. How might knowledge of these cache parameters be exploited in order to improve program performance? [5 marks]

What pros and cons might the designers have considered when selecting the cache line size? [5 marks]

The processor's L1 D-cache is indexed by virtual address and tagged by physical address. Why would the designers have done this? The processor has an 8 Kbyte page size, and so their decision is likely to have impact on the operating system virtual memory system. Explain why this is so, and state why the designers may have considered the possibility of increasing the cache's associativity. [5 marks]

The processor's physical address space is 64 bits wide. Calculate the number of bits of RAM required to implement the D-cache, including tag, data and status bits. [5 marks]

5 Business Studies

What is meant by a *critical path*? [5 marks]

Draw a PERT diagram for the process of getting up in the morning, making breakfast (tea, toast), and leaving for lectures. Show the critical path. [5 marks]

Derive the GANTT chart. What is the latest time to get up, if you must leave at 08:30? Comment on resource conflicts. [5 marks]

Does extra capital plant (such as an automatic tea-maker), allow you to get up later, and by how much? What other process improvements might there be? [5 marks]

6 Information Retrieval

Using natural language words directly for searching with a retrieval system has a number of disadvantages. What are they? [7 marks]

There are many different methods for compensating for the problems of simple word indexing. Describe *four* of these in detail, showing how they work and justifying their use in a system. [10 marks]

Which of these methods do you regard as most effective, and why? [3 marks]

7 Optimising Compilers

Consider a first-order call-by-need language with identifiers x_i and expressions e_i (which can be of type `int` only) and function names A_i (built-in) and G (a single, possibly recursive, user-defined function of the form $G(x_1, \ldots, x_k) = e$) whose arguments and results are of type `int`.

Describe the basic concepts of strictness analysis. You should explain what space of abstract values you would use to model strictness properties of a function of k arguments, and give the abstract strictness values for *cond* and *plus* (respectively the ternary conditional function and the binary addition function). State how one can determine that "f is strict in its i^{th} argument" in terms of your abstract value for a function f and how the abstract strictness value for G is obtained. [10 marks]

As an alternative method of deriving strictness properties, it is proposed to use an effect-like system instead. Suppose f is a function of k arguments and S is a subset of $\{1, \ldots, k\}$. In such a system judgements *on functions* f are of the form

$$\Gamma \vdash f : \texttt{int}^k \xrightarrow{S} \texttt{int}$$

where Γ is a set of type assumptions on variables x. The above judgement is defined to be valid if, whenever f is applied and all argument expressions in argument positions S fail to terminate, then the call to f fails to terminate (such f are often called S-jointly strict). In the following t ranges over type and effect forms $\texttt{int}^k \xrightarrow{S} \texttt{int}$.

Give an inference rule (here called (SUB)) for judgements of the form $\Gamma \vdash f : t$ which captures the idea that "if f is S-jointly strict in argument positions S, then a call to it fails to terminate when applied to argument expressions which fail to terminate for a larger set of argument positions". [3 marks]

Give a suitable set of assumptions of the form

$$\Gamma_0 = \{plus : t_1, plus : t_2, cond : t_3, cond : t_4\}$$

which together with the (SUB) rule above enable one to deduce exactly the valid strictness judgements $\Gamma_0 \vdash f : t$ when f is *plus* or *cond*. [Hint: there are two t_i for both *plus* and *cond*.] [4 marks]

Give conditions on t in the judgement $\Gamma \vdash f : t$ which enables the claim "f is strict in its i^{th} argument" to be made. [3 marks]

8 Artificial Intelligence

The following stories are adapted from P. Winston's textbook *Artificial Intelligence* (2nd edition, 1984):

Thomas and Albert

Thomas and Albert respected each other's technical judgement and decided to form a software company together. Unfortunately, Thomas learned that Albert was notoriously absentminded, whereupon he insisted that Albert have nothing to do with the proposed company's finances. This angered Albert so much that he backed out of their agreement, hoping Thomas would be disappointed.

John and Mary

John and Mary loved each other and decided to be married. A month before the wedding, John discovered that Mary's father was secretly smuggling stolen art through Venice. After struggling with his conscience for days, John reported Mary's father to the police. Mary understood John's decision, but she despised him for it nonetheless, and she broke off their engagement knowing he would suffer.

For each story, devise a semantic network to represent the characters and events in the story. [10 marks]

Show how a correspondence may be made between these semantic networks to represent the abstract similarity between these stories. [10 marks]

9 Neural Computing

Explain why probability theory plays a central role in neural computation. Discuss how the problem of classification can be expressed in terms of the estimation of a probability distribution. [6 marks]

Explain what is meant by a *likelihood function* and by the concept of *maximum likelihood*. [4 marks]

Consider a neural network regression model which takes a vector $\mathbf{x}$ of input values and produces a single output $y = f(\mathbf{x}, \mathbf{w})$ where $\mathbf{w}$ denotes the vector of all adjustable parameters ("weights") in the network. Suppose that the conditional probability distribution of the target variable t, given an input vector $\mathbf{x}$, is a Gaussian distribution of the form

$$p(t|\mathbf{x}, \mathbf{w}) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{\{t - f(\mathbf{x}, \mathbf{w})\}^2}{2\sigma^2}\right)$$

where σ^2 is the variance parameter.

Given a data set of input vectors $\{\mathbf{x_n}\}$, and corresponding target values $\{t_n\}$, where $n = 1, \ldots, N$, write down an expression for the likelihood function, assuming the data points are independent. Hence show that maximisation of the likelihood (with respect to $\mathbf{w}$) is equivalent to minimisation of a sum-of-squares error function. [10 marks]

10 Natural Language Processing

Consider the following dialogue between a user (U) and a natural language processing system (S).

U: There are three ships in the Pacific near the Johnson Islands and four in the Atlantic off Bermuda. The aircraft carrier, *Nimitz*, is in the Pacific with two frigate escorts, *Ulysses* and *Poseidon*. The *Enterprise*, which is a rapid assault ship, is in the Atlantic escorted by the frigate, *Medusah*, and two others.

S: OK

U: How many frigates are at sea?

S: 5

Describe one set of techniques that the system might utilise to support the dialogue. [8 marks]

What problems and challenges would the dialogue pose for these techniques? [12 marks]

11 Information Theory and Coding

What is the entropy H, in bits, of the following source alphabet whose letters have the probabilities shown?

A	B	C	D
1/4	1/8	1/2	1/8

[2 marks]

Why are fixed length codes inefficient for alphabets whose letters are not equiprobable? Discuss this in relation to Morse Code. [4 marks]

Offer an example of a uniquely decodable prefix code for the above alphabet which is optimally efficient. What features make it a uniquely decodable prefix code? [6 marks]

What is the coding rate R of your code? How do you know whether it is optimally efficient? [4 marks]

What is the maximum possible entropy H of an alphabet consisting of N different letters? In such a maximum entropy alphabet, what is the probability of its most likely letter? What is the probability of its least likely letter? [4 marks]

12 Computer Vision

It is often useful in computer vision to represent and analyse image content by means of complex variables, even though an image itself is defined as an array of real numbers. Give at least two distinct examples of useful operations in computer vision based on complex variables, identifying clearly the mathematical domain in which the complex variables exist. Explain in each case what is achieved by adopting such a representation. [10 marks]

In visual pattern recognition algorithms employing complex-valued wavelets, the twin tasks of classification and of discrimination among members of a class are handled differently. When the real and the imaginary parts of wavelet representations are resolved into their complex polar form as modulus and phase, what kind of information is extracted by the modulus? What kind by the phase? Use the examples of detecting faces and of identifying faces to illustrate your answer. [10 marks]

13 Specification and Verification II

Describe briefly how propositional boolean formulae can be represented as Binary Decision Diagrams (BDDs). [6 marks]

What is the significance of variable ordering? [2 marks]

Describe how the BDDs representing existentially and universally quantified boolean formulae are constructed. [2 marks]

Draw the BDDs of both $\neg(x = y) \Rightarrow z$ and $\forall z.\ \neg(x = y) \Rightarrow z$. [4 marks]

Let $E_1(x, y)$ and $E_2(x, y)$ be boolean formulae containing the variables x and y. Let the relation $\mathcal{R}$ be defined by:

$$\mathcal{R}((x, y), (x', y')) \ = \ (x' = E_1(x, y) \ \wedge \ y' = y) \ \vee \ (x' = x \ \wedge \ \ y' = E_2(x, y))$$

Assuming you have already computed the BDD of $P(a, b)$, explain how the BDD of $\exists a\ b.\ P(a, b) \ \wedge \ \mathcal{R}((a, b), (x, y))$ can be computed without having to compute the BDD of $\mathcal{R}((a, b), (x, y))$. Explain the significance of this. [6 marks]

14 Numerical Analysis II

State a recurrence formula for the sequence of Chebyshev polynomials, $\{T_n(x)\}$, and list these as far as $T_5(x)$. [4 marks]

What is the best polynomial approximation over $[-1, 1]$ to x^n using polynomials of lower degree, and what is its degree? Use this property to explain the method of economisation of a Taylor series. How can the error in one economisation step be estimated? [7 marks]

The error in Lagrange interpolation can be expressed in the form

$$f(x) - L_{n-1}(x) = \frac{f^n(\xi)}{n!} \prod_{j=1}^{n} (x - x_j)$$

for a suitable function $f(x)$. What is the best choice for abscissae $\{x_j\}$ and why? [2 marks]

The function $\sin x$ may be approximated by the truncated Taylor series

$$P_{2n-1}(x) = \sum_{i=1}^{n} (-1)^{i-1} \frac{x^{2i-1}}{(2i-1)!}.$$

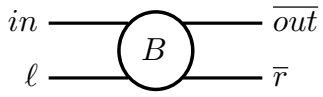
Estimate the maximum absolute error over $[-1, 1]$ for both $P_3(x)$ and $P_5(x)$. Perform one economisation step on $P_5(x)$ and show that the resulting polynomial is more accurate than $P_3(x)$. [7 marks]

15 Communicating Automata and Pi Calculus

Explain the notions of *abstraction* and *concretion* in the π-calculus. Explain the components of a *commitment* $P \xrightarrow{\alpha} A$, and say what it means for each form which α may take. (You need not give the rules of commitment.) Define *strong bisimulation* in terms of commitments. [5 marks]

Consider each pair of the three processes $(\mathsf{new}\, x)\overline{x}\langle y\rangle$, $(\mathsf{new}\, x)\overline{y}\langle x\rangle$, and $\mathsf{0}$. Are they structurally congruent ($\equiv$)? Are they strongly equivalent ($\sim$)? Briefly justify each of your six answers. [4 marks]

The following equations define the behaviour of a buffer cell which has the ability to cut itself out of a chain of similar cells:



$$B(in, out, \ell, r) \stackrel{\text{def}}{=} in(x).C\langle x, in, out, \ell, r\rangle + \overline{r}\langle in, \ell\rangle.\mathsf{0}$$

$$C(x, in, out, \ell, r) \stackrel{\text{def}}{=} \overline{out}\langle x\rangle.B\langle in, out, \ell, r\rangle + \ell(in', \ell').C\langle x, in', out, \ell', r\rangle$$

Let $P = \mathsf{new}\, mid\, m\, (B\langle in, mid, \ell, m\rangle \mid C\langle x, mid, out, m, r\rangle)$. Express P as a summation up to $\sim$, i.e. $P \sim \Sigma\alpha_i A_i$. Use structural congruence to make the expression as simple as possible. Justify your expression. [6 marks]

Now suppose that the name *out* is replaced by *in* in the definition of P. What effect does this have upon the behaviour of P? Briefly justify your answer in terms of commitments. [5 marks]