COMPUTER SCIENCE TRIPOS Part II

Wednesday 4 June 1997 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

Outline the main points of *each* of the following topics, using examples where appropriate:

(*a*) why different WHILE-rules are needed for partial and total correctness

(*b*) verification conditions

(*c*) Floyd–Hoare logic as higher-order logic

(*d*) soundness and completeness for Floyd–Hoare logic

[5 marks each]

1   **[TURN OVER**

2 Digital Communication II

In the Internet, error-detecting check fields are usually included at the data-link, network and transport layers of the protocol stack. Why is this the case? [6 marks]

What is the standard Internet checksum? [3 marks]

Explain in detail how the Internet checksum can be efficiently implemented, commenting on how the word size and machine endian-ness affect your implementation. [5 marks]

A certain engineer constructs an Ethernet interface which includes processing at the IP layer in hardware. The engineer "tests" the new interface by sending many 1 kbyte UDP datagrams filled with different random contents from a conventional workstation to a host with the new interface, and observes that the UDP checksum software on the host detects no errors. In fact, the engineer has incorrectly used inverting pads on those pins of the ASIC used to transfer the UDP data and UDP header from the interface to the host memory. The IP header is transferred over other (non-inverting) pins.

How can it be that no errors were detected? What can be deduced about the two hosts being used? [6 marks]

[Hint: you may wish to consider the rôle of the UDP pseudo-header.]

3 Computer System Modelling

A group of smart computer scientists decide to earn some money in the vacation by buying a punt and running a chauffeured punt tour operation on the Cam. Based at Magdalene Bridge, they run tours upriver showing tourists the delights of the Backs. As those familiar with punt trips will know, journeys take an exponentially distributed time to complete, particularly when accompanied by strawberries and champagne. In this case our punt operators have measured that the average trip takes 30 minutes. Tourists (being tourists) behave rather randomly and arrive at the quay independently, according to Poisson distribution with mean 10 per hour. Each punt can accommodate six tourists, but departs as soon as there is at least one tourist wanting to take the tour.

(a) Draw a Markov chain model for the punt tour operation, annotating it with the appropriate rates and state information. Briefly explain the diagram.

[5 marks]

(b) Propose a model which can safely be used as a worst-case approximation to the system above. Why is your model conservative? Using your new model, calculate the expected length of the queue of tourists. What constraints are required to ensure that the queuing system remains stable?   [5 marks]

(c) Now assume that tourists do not join the queue if there are already six others waiting. If each tourist pays £5 per trip, and a punt costs £10 per hour to run, calculate the expected hourly profit of the punt company. Propose a scheme which would enable the entrepreneurs to increase their profits, but do not solve a model of the new system. What is the utilisation of the punts in the old and new schemes?   [10 marks]

4   Comparative Architectures

A number of techniques exist to reduce the penalty of control-flow hazards. Discuss *four* of the following:

(*a*) delayed branches and annulment

(*b*) predicated execution and conditional moves

(*c*) dynamic branch prediction caches

(*d*) jump target hints and subroutine call return prediction stacks

(*e*) next-fetch predictors and branch target buffers

[4 marks each]

Why is good control-flow prediction so important to a dynamically scheduled superscalar processor?   [4 marks]

5   Business Studies

What is meant by a project's *critical path* and why is it important?   [5 marks]

An organisation which manufactures about 30 different kinds of widget has asked you to develop a web site which includes an on-line catalogue and order entry facility. You may assume that service provision (including access to a merchant system) will be provided by a service provider, at no cost to this project.

Draw up a task list and project plan, and indicate the critical path.   [5 marks]

Estimate a price and time to completion for the job, and how much working capital will be required, stating your assumptions.   [5 marks]

What precautions can be taken to ensure that the project completes to the customer's satisfaction?   [5 marks]

6 Advanced Algorithms

A posting on a newsgroup announces the invention of a new compression algorithm, and claims that the method will guarantee to compress at least 10% of all possible input files by at least 10% of their original size, but that it might (unfortunately) cause some of the other 90% of possible inputs to expand by a factor of up to 90.

Discuss how believable and reasonable the claim is.   [8 marks]

Various compression utilities running on personal computers typically reduce text files to 1/3 of their original size. Estimate the proportion of all possible files (including binary ones) whose original length is less than 64 kbytes that could be compressed to this extent by an ideal compression algorithm. What proportion of the same set of files consists of just alphanumeric characters, blanks and newlines?
                    [12 marks]

7 Optimising Compilers

Summarise the idea of the 3-address instruction and briefly indicate its advantage over stack-oriented instructions as intermediate code. Explain the notion of the flowgraph containing such instructions, giving such a flowgraph for the C function:

```
int f(int x, int g())
{   return x==0 ? g(x) : x*f(x-1, g);
}
```

You should explain how arguments and results of functions in your example are communicated; also make clear any difference in the representation of calls to f and g. [8 marks]

Given a flowgraph in which each node contains a single 3-address instruction (not grouped into basic blocks), design dataflow equations (and thence an algorithm) for *reaching definitions*. [9 marks]

The reaching definitions $RD(n)$ of a node n are the set of nodes m such that m contains a 3-address instruction such as

$$m : x := a + b$$

which writes to ("defines") one of its operands (here x), and such that there is a path in the flowgraph from m to n by which the value given to x at m *may* still be unchanged (by other assignments to x) when it reaches n. Hence in

```
1: y := 1;
2: if x<=1 goto 5;
3: y := x*y;
4: x := x-1; goto 2;
5: return
```

we would have $RD(3) = \{1, 3, 4\}$ and $RD(4) = \{3, 4\}$; note that the definition (of y) at 3 reaches via the loop back to 3 even though it would not be *available* because of node 4.

Develop a program in which $m \in RD(n)$ but which no run-time execution can cause the value assigned at m to reach n. To what extent can we fix this problem? [3 marks]

[Hint: work by analogy from live variable or available expression analysis; determine the direction of the analysis and suitable *gen* and *kill* properties of nodes. You may find it convenient to consider cases like "l contains a statement $x := e$" or "l contains some other statement".]

6

8   Artificial Intelligence

Describe the Minimax Algorithm for searching game trees.   [5 marks]

Explain how the Alpha-Beta Algorithm is a better way to search game trees.
[10 marks]

These two algorithms depend on certain assumptions about how the game is played.
What are they?   [5 marks]

9   Database Topics

Describe briefly what is involved in extending a programming language to support
persistence of data.   [5 marks]

What are the advantages and drawbacks of using a persistent programming
language for database management as opposed to a conventional DBMS based
on the ANSI/SPARC architecture?   [8 marks]

To what extent may it be said that the ODMG proposals offer the advantages of
persistent programming within a framework of conventional database management?
[7 marks]

10   Information Retrieval

Give a formula for weighting the index terms in a query for an initial search of
a document file, and indicate how an overall query-document matching score is
computed.   [5 marks]

Outline the main features of the general model underpinning this method of
weighting and scoring.   [5 marks]

What is the motivation for using the various types of information employed in the
method you have described?   [5 marks]

What further information could you use after conducting an initial search over the
file? Sketch how you would apply this to modify the original query and construct
a new one.   [5 marks]

11 Information Theory and Coding

The input source to a noisy communication channel is a random variable X over the four symbols a, b, c, d. The output from this channel is a random variable Y over these same four symbols. The joint distribution of these two random variables is as follows:

	$x = a$	$x = b$	$x = c$	$x = d$
$y = a$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$y = b$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	0
$y = c$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{16}$	0
$y = d$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{16}$	0

(a) Write down the marginal distribution for X and compute the marginal entropy $H(X)$ in bits. [3 marks]

(b) Write down the marginal distribution for Y and compute the marginal entropy $H(Y)$ in bits. [3 marks]

(c) What is the joint entropy $H(X, Y)$ of the two random variables in bits? [4 marks]

(d) What is the conditional entropy $H(Y|X)$ in bits? [4 marks]

(e) What is the mutual information $I(X; Y)$ between the two random variables in bits? [2 marks]

(f) Provide a lower bound estimate of the channel capacity C for this channel in bits. [2 marks]

(g) Draw a Venn diagram that describes the relationships among the quantities $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, $H(X, Y)$, and $I(X; Y)$. [2 marks]

12   Computer Vision

Explain the notion of scale-space and how it is used in various areas of computer vision. Include the following:

(*a*) Pyramidal representations of image structure across successive scales of blurred undersampling.   [5 marks]

(*b*) Edge detection operators that extract edges at particular scales of analysis, but not at others.   [5 marks]

(*c*) The behaviour of zero-crossings, their trajectories and "fingerprints" in scale-space.   [5 marks]

(*d*) The generalised wavelet transform as a self-similar mapping into scale-space, and its attempt to capture invariances under the transformations of dilation, translation and rotation.   [5 marks]

13   Specification and Verification II

In the context of the specification and verification of hardware, explain the key ideas of *each* of the following:

(*a*) binary decision diagrams

(*b*) temporal logic

(*c*) model checking

(*d*) higher-order predicates

[5 marks each]

14   Numerical Analysis II

Define the *Chebyshev polynomial* $T_k(x)$.   Evaluate $T_4(\frac{1}{2})$ using the formula $T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x)$. What is the leading coefficient of $T_k(x)$?   [4 marks]

The best L_∞ approximation to $f(x) \in C[-1,1]$ by a polynomial $p_{n-1}(x)$ of degree $n-1$ has the property that
$$\max_{x\in[-1,1]} |e(x)|$$
is attained at $n+1$ distinct points $-1 \leqslant \xi_0 < \xi_1 < \ldots < \xi_n \leqslant 1$ such that $e(\xi_j) = -e(\xi_{j-1})$ for $j = 1, 2, \ldots n$.

Let $f(x) = x^2$. Show, by means of a clearly labelled sketch graph, that the best polynomial approximation of degree 1 is a constant.   [3 marks]

Now suppose $f(x) = x^3$ is the function to be approximated. Taking account of symmetry, sketch the graph of $f(x)$ and its best L_∞ approximation by a polynomial of degree 2.   [5 marks]

By differentiating $e(x)$, find the polynomial $p_2(x)$.   [6 marks]

State a formula for the best approximation to $f(x) = x^n$ by a polynomial of degree $n-1$.   [2 marks]

15   Communicating Automata and Pi Calculus

The main reaction rule of the π-calculus is

$$COMM : (M + x(y).P) \mid (N + \overline{x}\langle z\rangle.Q) \longrightarrow \{z/y\}P \mid Q$$

What other reaction rules are needed to infer all reactions?   [4 marks]

Using these rules, show how to infer a reaction on the x-channel for the following process:
$$(!x(y).\overline{z}\langle y\rangle) \mid (\boldsymbol{\nu}\, z)\big(Q \mid (\overline{x}\langle z\rangle.R + \overline{y}\langle z\rangle.S)\big)$$

Indicate exactly which rules of structural congruence are required in making the inference.   [7 marks]

Let $P \longrightarrow P'$ be an arbitrary reaction inferred for a process P not containing "+". Prove, by induction on the depth of the inference, that if P contains an instance of replication "!" then so does P'.   [6 marks]

Show that this is not necessarily true if P contains "+".   [3 marks]