COMPUTER SCIENCE TRIPOS Part II

Wednesday 7 June 2006 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.*

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

STATIONERY REQUIREMENTS
Script Paper
Blue Coversheets
Tags

1 Comparative Architectures

(*a*) One or more processors must be selected for inclusion in a portable device that handles audio, image, video, telephone calls and simple word processing.

Possible design combinations include:

- a single processor;
- a pair of basically similar processors but with different coprocessor and bus structures;
- a standard processor and a custom VLIW processor.

Explain whether one of these combinations is clearly the best. [2 marks]

(*b*) For *each* of the following processor technologies, justify whether it is applicable in the device:

(*i*) Dynamically Scheduled Instruction Dispatch;

(*ii*) Simultaneous Multi-Threading (SMT);

(*iii*) SIMD extensions, similar to Intel's MMX;

(*iv*) Virtual Memory;

(*v*) Dynamic Clock Frequency/Power Control;

(*vi*) Snooping Cache.

[3 marks each]

2 VLSI Design

A *majority gate* with three inputs signals logic one on its output if two or more of its inputs are one, and zero otherwise. A *minority gate* is its complement.

(*a*) Give the boolean equation for a minority gate as a sum of products. [2 marks]

(*b*) Sketch the circuit diagram for a minority gate using NAND gates and inverters. [3 marks]

(*c*) Sketch the transistor-level circuit diagram for an alternative implementation as a single stage of CMOS logic. [3 marks]

(*d*) Calculate the number of transistors required for each implementation. [2 marks]

(*e*) Assuming that a conducting p-channel has a resistance twice that of a similarly sized n-channel, use logical effort to compare the performance of the two implementations. [10 marks]

3 Digital Communication II

(*a*) The Transmission Control Protocol (TCP) employs a transmit window and cumulative acknowledgement and timeout system, as well as sequence numbering of packets, to achieve reliable delivery of data.

(*i*) Outline the procedure for round-trip time estimation and the calculation of the retransmission timer. [8 marks]

(*ii*) Explain the function of buffering at the sender and receiver. [3 marks]

(*iii*) How do fast retransmit and fast recovery improve performance after packet loss? [3 marks]

(*b*) In a wireless network, delay to access the channel due to scheduling of the media access control protocol, and random packet loss due to interference, may be non-negligible.

(*i*) Explain how this can interfere with the round-trip time estimation process above. [3 marks]

(*ii*) Explain how this can interfere with the congestion control scheme that TCP employs. [3 marks]

4 Distributed Systems

(*a*) Describe, with examples, the function of a naming service for a large-scale distributed system. Include definitions for "name space" and "naming domain". [6 marks]

(*b*) Discuss consistency *versus* availability for naming data in large-scale systems. [4 marks]

(*c*) How can any distributed naming service be engineered so that invocations on behalf of users can be resolved efficiently in the presence of failures and heavy load? [4 marks]

(*d*) Contrast the assumptions under which DNS was designed originally for the Internet, with the properties of dynamically formed groups of mobile hosts using wireless communication (MANETS). How might DNS-like services be provided for MANETS? [6 marks]

5 Advanced Systems Topics

Modern peer-to-peer (P2P) systems are typically described as *structured* or *unstructured*.

(*a*) Compare and contrast these two approaches. Include a discussion of the general topology, membership management and query mechanisms. Use examples to support your answer. [6 marks]

(*b*) Which approach is more resilient to *churn*? Justify your answer. [2 marks]

(*c*) One early criticism of P2P systems was that they did not consider network latencies. Describe how one can add proximity awareness to:

(*i*) unstructured P2P systems; [1 mark]

(*ii*) structured P2P systems. [3 marks]

(*d*) *Swarming* P2P systems like BitTorrent are designed for efficient (and incentive compatible) download of large files. However, it is typically not possible to use the file until it has downloaded in its entirety. Sketch the design of a swarming P2P system which supports streaming video – that is, allows playback of video to overlap the ongoing download of the remainder of the stream. Comment on how efficient (in terms of network resources) your system would be in comparison with a system like BitTorrent. [8 marks]

6 Computer Vision

(*a*) Explain the method of *Active Contours.* What are they used for, and how do they work? What underlying trade-off governs the solutions they generate? How is that trade-off controlled? What mathematical methods are deployed in the computational implementation of Active Contours? [10 marks]

(*b*) When trying to detect and estimate visual motion in a scene, why is it useful to relate spatial derivatives to temporal derivatives of the image data? Briefly describe how one motion model works by these principles. [5 marks]

(*c*) Provide a 3×3 discrete filter kernel array that approximates the Laplacian operator. Explain what the Laplacian might be used for, and what is the significance of the sum of all of the taps in the filter. [3 marks]

(*d*) When visual sequences are encoded into an `.mpeg` video stream, typically about what percentage of the compression achieved is intra-frame (compression within individual still frames), and what percentage is inter-frame? Name a key feature that is extracted and estimated for purposes of prediction and, therefore, compression. [2 marks]

7 Security

The Needham–Schroeder protocol is defined as

1. $A \longrightarrow S : A, B, N_A$
2. $S \longrightarrow A : \{N_A, B, K_{AB}, \{K_{AB}, A\}_{K_{BS}}\}_{K_{AS}}$
3. $A \longrightarrow B : \{K_{AB}, A\}_{K_{BS}}$
4. $B \longrightarrow A : \{N_B\}_{K_{AB}}$
5. $A \longrightarrow B : \{N_B - 1\}_{K_{AB}}$

(*a*) Explain the symbolism, and the purpose of the messages. [5 marks]

(*b*) Explain the "bug" in the protocol. [5 marks]

(*c*) Is the bug actually a vulnerability if one can assume (as the Needham–Schroeder paper does) that all principals execute the protocol faithfully? If not, why is it important? [5 marks]

(*d*) Describe how *one* modern protocol derived from Needham–Schroeder deals with the issue. [5 marks]

8 Optimising Compilers

(*a*) Summarise the idea of a *basic block* and explain why it is useful in intermediate representations for optimising compilers. [3 marks]

(*b*) Construct the flowgraph (in which every node is a basic block consisting of one or more 3-address instructions) for the C function:

```
int f(int x, int y)
{
    int r = x + 1;
    if (y == 0) {
        r = r * r;
    } else {
        y = y - 1;
        r = r * y;
    }
    return r + 1;
}
```

[4 marks]

(*c*) Define *static single assignment* (SSA) form, and explain the changes you would have to make to your flowgraph from part (*b*) in order for it to be in SSA form. [3 marks]

(*d*) Consider a flowgraph in which every node contains a single 3-address instruction. Each node whose instruction assigns some value to a variable is considered a "definition" of that variable; we are interested in discovering, for each node n in the flowgraph, which definitions *reach* n. A definition m is considered to reach n if the variable to which m assigns a value may still have that value at entry to n.

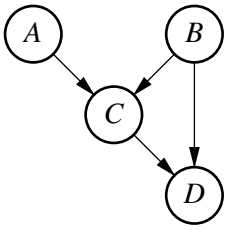
(*i*) Define the notion of a definition reaching a node in the flowgraph in terms of possible execution flows of control. [2 marks]

(*ii*) By analogy with live variable or available expression analysis, or otherwise, design dataflow equations for computing $RD(n)$, the set of definitions which can reach a node n. [4 marks]

(*iii*) Sketch an algorithm to compute $RD(n)$, briefly commenting on any initialisation required. [4 marks]

9 Artificial Intelligence II

Consider the following Bayesian network:



The associated probability distributions for the binary random variables A, B, C and D are $\Pr(a) = 0.1$, $\Pr(\neg a) = 0.9$, $\Pr(b) = 0.8$, $\Pr(\neg b) = 0.2$, and:

A	B	$\Pr(c\|A,B)$
$\top$	$\top$	0.5
$\top$	$\bot$	0.6
$\bot$	$\top$	0.8
$\bot$	$\bot$	0.7

B	C	$\Pr(d\|B,C)$
$\top$	$\top$	0.2
$\top$	$\bot$	0.9
$\bot$	$\top$	0.8
$\bot$	$\bot$	0.1

(a) Explain why the representation of the joint distribution of A, B, C and D using the Bayesian network is preferable to a direct tabular representation. [2 marks]

(b) Use the *variable elimination algorithm* to compute the probability distribution of B conditional on the evidence that $D = \top$. [16 marks]

(c) Comment on the computational complexity of the variable elimination algorithm. [2 marks]

10 Digital Signal Processing

(a) Consider a software routine that converts and records the audio samples received in a digital telephone network call (8 kHz sampling frequency, 8 bit/sample) into a WAV file (8 kHz sampling frequency, 16 bit/sample, uniform quantisation). Your colleague attempted to write a very simple conversion routine for this task, but the resulting audio is very distorted.

(i) Name *two* variants of the method used for quantising the amplitude of audio samples in digital telephone networks and explain *one* of them. [4 marks]

(ii) Your colleague's routine right-pads each 8-bit data word from the telephone network with eight additional least-significant zero bits to obtain 16-bit values. Explain how this distorts the signal by discussing which frequencies could appear at the output when the incoming telephone signal consists of a pure 1 kHz sine tone. [4 marks]

(b) A real-valued discrete random sequence $\{x_i\}$ is fed into a linear time-invariant filter with impulse response $h_0 = 1$, $h_3 = 1$, and $h_i = 0$ for all other i. We observe for the resulting output sequence $\{y_i\}$ the expected value

$$\mathcal{E}(y_{i+k} \cdot x_i) = \begin{cases} 1 & \text{for } k = -1 \\ 2 & \text{for } k = 0 \\ 1 & \text{for } k = 1 \\ 1 & \text{for } k = 2 \\ 2 & \text{for } k = 3 \\ 1 & \text{for } k = 4 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of the autocorrelation sequence $\{\phi_{xx}(k)\}$? [4 marks]

(c) The $YCrCb$ colour encoding is used in many image compression methods.

(i) How is it defined and why is it used? [4 marks]

(ii) Is the conversion from 3×8-bit RGB to 3×8-bit $YCrCb$ coordinates fully reversible? Why? [4 marks]

11 Computer Systems Modelling

(*a*) Define the $M/M/1$ queueing model and derive the steady-state distribution for the number of customers present when the traffic intensity is less than one. [5 marks]

(*b*) For the $M/M/1$ model in steady-state, derive the mean number of customers present and the mean time spent by a customer in the system. What is the utilisation of the server? [5 marks]

(*c*) Now consider the $M/M/1/K$ queueing model with K finite and again derive the steady-state distribution for the number of customers present. For what values of the traffic intensity does your steady-state distribution exist? What is the utilisation of the server and explain how this compares to the $M/M/1$ queueing model. [5 marks]

(*d*) Give an example of the use of the $M/M/m/m$ loss model. Derive Erlang's formula for the steady-state probability that an arriving customer finds all m servers occupied. [5 marks]

12 Numerical Analysis II

(a) In Peano's theorem, if a quadrature rule integrates polynomials of degree N exactly over an interval $[a, b]$, then the error in integrating $f \in C^{N+1}[a, b]$ is 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, $(x-t)_+^N$. [4 marks]

(b) Assuming $x \in [a, b]$, and writing Taylor's theorem in the form

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

where P_N is a polynomial of degree N, prove Peano's theorem, explaining each step clearly. [8 marks]

(c) For the trapezium rule, what is N? [1 mark]

(d) If $K(t)$ does not change sign in $[a, b]$ then

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

for some $\xi \in (a, b)$. Use this result to simplify

$$E(f) = \int_{-1}^{1} f(x)\, dx - f(-1) - f(1).$$

[7 marks]

13 Bioinformatics

(*a*) Why do we use dynamic programming algorithms for pairwise sequence alignment problems but not for multiple pairwise alignment? [5 marks]

(*b*) Compare the use of the affine gap penalty with the constant gap penalty. [3 marks]

(*c*) Discuss the properties and assumptions of the Jukes–Cantor and the Kimura 2-parameter models of DNA evolution. [5 marks]

(*d*) Describe the UPGMA algorithm. [4 marks]

(*e*) What does the ultrametric property of a tree tell us about the evolutionary process? [3 marks]

14 Denotational Semantics

Let D be a domain with bottom element $\bot$. Let $h, k : D \to D$ be continuous functions with h strict (so $h(\bot) = \bot$). Let $\mathbb{B} = \{\mathit{true}, \mathit{false}\}$. Define the conditional function

$$\mathit{if} : \mathbb{B}_\bot \times D \times D \to D$$

by $\mathit{if}(b, d, d') = d$ if $b = \mathit{true}$, d' if $b = \mathit{false}$, and $\bot$ otherwise. Let $p : D \to \mathbb{B}_\bot$ be a continuous function.

The function f is the least continuous function from $D \times D$ to D such that

$$\forall x \in D.\ f(x, y) = \mathit{if}(p(x),\ y,\ h(f(k(x), y))) \ .$$

(a) State the principle of fixed-point induction. What does it mean for a property to be chain closed? [4 marks]

(b) Assume that the property

$$Q(g) \Leftrightarrow_{\mathit{def}} \forall x, y \in D.\ h(g(x, y)) = g(x, h(y)) \ ,$$

where g is a continuous function from $D \times D$ to D, is chain closed. Prove $Q(f)$ by fixed-point induction. [7 marks]

(c) Let g be a continuous function from a cpo D to a cpo E. Let Y be a chain-closed subset of E. Show that the inverse image $g^{-1}Y$ is a chain-closed subset of D. [4 marks]

(d) Exhibit cpos D, E, F and chain-closed subsets $R \subseteq D \times E$ and $S \subseteq E \times F$ such that their relational composition $S \circ R \subseteq D \times F$ is not chain closed. (No proof is required.) [5 marks]

15 Specification and Verification I

If C is a command that contains one or more occurrences of a command `BREAK`, then `LOOP (`C`)` is a command that repeatedly executes C until a `BREAK` is executed. Executing `BREAK` immediately terminates the execution of `LOOP (`C`)`.

How might ideas from Floyd–Hoare Logic be used to verify programs that use the `LOOP/BREAK` construct, such as the following program that sets the variable `RES` to the factorial of the initial value of the variable `X` (if `X > 0`)?

```
RES := 1;
LOOP (IF X=1 THEN BREAK ELSE RES := RES×X; X := X-1)
```

You may place restrictions that you consider reasonable on the form of C, but these should be discussed and motivated.

[20 marks]

16 Information Theory and Coding

(*a*) Suppose we know the conditional entropy $H(X|Y)$ for two slightly correlated discrete random variables X and Y. We wish to guess the value of X, from knowledge of Y. There are $\mathcal{N}$ possible values of X. Give a lower bound estimate for the probability of error, when guessing X from knowledge of Y. What is the name of this relationship? [4 marks]

(*b*) In an error-correcting (7/4) Hamming code, under what circumstance is there still a residual error rate? (In other words, what event causes this error-correction scheme to fail?) [2 marks]

(*c*) Broadband noise whose power spectrum is flat is "white noise". If the average power level of a white noise source is σ^2 and its excursions are zero-centred so its mean value is $\mu = 0$, give an expression describing the probability density function $p(x)$ for excursions x of this noise around its mean, in terms of σ. What is the special relationship between the entropy of a white noise source, and its power level σ^2? [4 marks]

(*d*) Explain the phenomenon of aliasing when a continuous signal whose total bandwidth extends to $\pm W$ is sampled at a rate of $f_s < 2W$. If it is not possible to increase the sampling rate f_s, what can be done to the signal before sampling it that would prevent aliasing? [5 marks]

(*e*) Prove that the sinc function,

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

is invariant under convolution with itself: in other words that the convolution of a sinc function with itself is just another sinc function. You might find it useful to recall that the Fourier transform of a sinc function is the rectangular pulse function:

$$\Pi(k) = \begin{cases} \frac{1}{2\pi} & |k| \leq \pi \\ 0 & |k| > \pi \end{cases}$$

[5 marks]

END OF PAPER