COMPUTER SCIENCE TRIPOS Part II

Wednesday 3 June 2009 1.30 to 4.30

COMPUTER SCIENCE 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.*

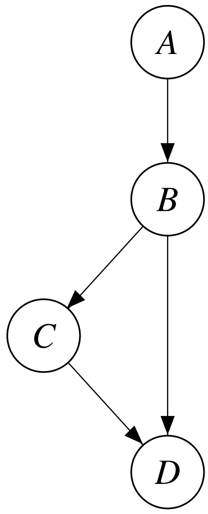
You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

STATIONERY REQUIREMENTS
Script paper
Blue cover sheets
Tags

SPECIAL REQUIREMENTS
Approved calculator permitted

1 Artificial Intelligence II

Consider the following Bayesian Network:



The associated probability distributions for the binary random variables A, B, C and D are $\Pr(a) = 0.7$, $\Pr(\neg a) = 0.3$ and:

A	$\Pr(b\|A)$	B	$\Pr(c\|B)$	B	C	$\Pr(d\|B, C)$
$\top$	0.1	$\top$	0.2	$\top$	$\top$	0.6
$\bot$	0.15	$\bot$	0.95	$\top$	$\bot$	0.5
				$\bot$	$\top$	0.4
				$\bot$	$\bot$	0.3

(a) Write down an expression for the full joint distribution of the random variables A, B, C and D. Compute the probability that A and B are $\top$ while C and D are $\bot$. [2 marks]

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

(c) Explain why the variable elimination might not be an effective algorithm to use in practice and suggest an alternative that addresses the shortcoming you have given. [2 marks]

2 Comparative Architectures

(*a*) How does the use of condition codes (also known as condition bits or status flags) complicate the implementation of a superscalar processor that supports out-of-order execution? [4 marks]

(*b*) A branch predictor with a high prediction accuracy is often employed to enable deeply pipelined processors to be exploited. What limits the complexity and size of such a branch predictor? [4 marks]

(*c*) In the best case, how can a branch predictor and branch target buffer enable a branch instruction and the instruction at the branch's target address to be fetched in consecutive clock cycles? [6 marks]

(*d*) Loop unrolling and predicated execution are two techniques that may be used to improve the performance of loops.

(*i*) How do these techniques improve performance? [3 marks]

(*ii*) What costs or disadvantages are associated with each technique? [3 marks]

3 Computer Systems Modelling

(*a*) Suppose that X is a random variable having the Binomial distribution with parameters n and p and that $\lambda > 0$ is a constant.

(*i*) Write down the expression for $\mathbb{P}(X = k)$ where $k \in \{0, 1, 2, \dots, n\}$ [2 marks]

(*ii*) Now suppose that $n \to \infty$ and p is chosen so that $p = \lambda/n$. Show that under this limit $\mathbb{P}(X = k) \to e^{-\lambda}\lambda^k/k!$, that is, to a Poisson distribution with parameter λ. [4 marks]

(*b*) Suppose that $N(t)$ is the random number of events in the time interval $[0, t]$ of a Poisson process with parameter λ.

(*i*) State the conditions that define the Poisson process $N(t)$. [2 marks]

(*ii*) Show that for all $t > 0$ the random variable $N(t)$ has the Poisson distribution with parameter λt. [4 marks]

(*c*) Given a Poisson process of rate λ let X_1 be the time of the first event and for $n > 1$ let X_n denote the time between the events $(n-1)$ and n. Thus the sequence $X_1, X_2, \dots$ gives us the sequence of *inter-event times* between the events in a Poisson process.

(*i*) Show that

$$\mathbb{P}(X_1 > t) = \mathbb{P}(N(t) = 0)$$

for $t > 0$. [4 marks]

(*ii*) Show that the inter-event times $X_1, X_2, \dots$ are independent, identically distributed random variables each of whose marginal distribution is an Exponential with rate parameter λ. [4 marks]

4 Computer Vision

(*a*) Discuss the significance of the fact that mammalian brains send almost ten times as many neural fibres back down the visual pathway (from cortex to thalamus), as there are ascending neural fibres bringing visual data from the retina up to the thalamus and cortex. Does this massive neural feedback projection support the thesis of "vision as hypothesis testing," and if so, how? Try to marshall other evidence supporting the view that in human vision "what you see is your own graphics" rather than the retinal image as faithfully recorded by photoreceptors in the eye. [5 marks]

(*b*) Write a block of pseudo-code for convolving an image with a feature-detecting kernel. (You may ignore out-of-bounds issues at the image array boundaries.) [3 marks]

(*c*) The Differentiation Theorem of Fourier analysis asserts that, for an image $f(x, y)$ whose 2D Fourier Transform (*2DFT*) is $F(\mu, \nu)$, taking derivatives of various orders m, n has the following spectral consequence:

$$\left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial y}\right)^n f(x,y) \stackrel{2DFT}{\Longrightarrow} (i\mu)^m (i\nu)^n F(\mu,\nu)$$

In particular when the 2nd-derivative Laplacian operator ∇^2 is applied to an image, its action is to multiply the Fourier transform of the image by a paraboloid:

$$\nabla^2 f(x,y) \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(x,y) \stackrel{2DFT}{\Longrightarrow} -(\mu^2 + \nu^2) F(\mu,\nu)$$

What therefore best describes the filtering operation that corresponds to taking derivatives of an image: lowpass, bandpass, or highpass? If noise in an image resides mainly in the higher spatial frequencies, should you use higher or lower order derivatives for edge detection to be less sensitive to the noise? What else could you do with differentiating filters to reduce their noise sensitivity? [3 marks]

(*d*) In pattern classification with two classes, explain how an ROC curve is derived from the underlying distributions. Define a threshold-independent performance metric based on the distributions' moments. [4 marks]

(*e*) When visually inferring a 3D representation of a face, it is useful to extract separately both a shape model, and a texture model. Explain the purposes of these steps, their use in morphable models for pose-invariant face recognition, and how the shape and texture models are extracted and later re-combined. [5 marks]

5 Digital Communication II

(*a*) Shared media networks make use of Media Access Control protocols to allocate resources.

(*i*) Outline the use of CSMA/CD on Ethernets. [5 marks]

(*ii*) Outline the use of CSMA/CA on Wireless Ethernets. [5 marks]

(*b*) Discuss the trade-offs between network resource pre-allocation and network resource contention-based schemes in terms of delays and throughput under varying traffic loads. [10 marks]

6 Digital Signal Processing

While reverse-engineering a radio receiver, you find in its firmware the following two discrete systems implemented:

$$y_n := x_n e^{j\pi n/2}$$

$$z_n := \sum_{k=-4000}^{4000} y_{n-k-4000} \times 10^{-3} \operatorname{sinc}(k/10^3) \times \left(0.54 - 0.46 \cos\left(2\pi \frac{k+4000}{8000}\right)\right)$$

The discrete sequence $\{x_n\}$ emerges from an analog-to-digital converter operating at sampling frequency $f_\mathrm{s} = 240$ kHz, whose input is connected via a 100 kHz low-pass filter and linear amplifier directly to a radio antenna.

(*a*) Explain the function of *both* discrete systems in the frequency domain and their main parameters (e.g. type of filter, cutoff frequency, type of window). [12 marks]

(*b*) In approximately which frequency range will antenna signals substantially influence the resulting sequence $\{z_n\}$? [4 marks]

(*c*) Will the subsequent application of the discrete system

$$b_n := z_{500n}$$

cause aliasing, and why? [4 marks]

7 Types

(*a*) Explain what is meant by a *solution* for a Mini-ML typing problem $\Gamma \vdash M : ?$ and what it means for a solution to be *principal*. [4 marks]

(*b*) Consider the following typing problems (where α and β are distinct type variables).

(*i*) $x : \forall\{\beta\}(\beta \to \alpha) \vdash x\,(x\,\mathtt{nil}) : ?$

(*ii*) $x : \forall\{\alpha\}(\beta \to \alpha) \vdash x\,(x\,\mathtt{nil}) : ?$

(*iii*) $x : \forall\{\beta\}(\beta \to \alpha\ \mathit{list}) \vdash x :: (x\,\mathtt{nil}) : ?$

(*iv*) $x : \forall\{\alpha\}(\beta \to \alpha\ \mathit{list}) \vdash x :: (x\,\mathtt{nil}) : ?$

For each typing problem, either give a solution together with a proof of typing, or show that no solution exists. [16 marks]

8 E-Commerce

It is said that the Internet played a significant role in the US presidential election. You are asked by a major UK political party to advise them on their Internet policy and associated website. The party has about 200,000 members, each paying a fee of £25 per annum. Its total annual income, mainly from donations, is about £20M. The UK has about 46 million people registered to vote. There are about 646 parliamentary constituencies.

(*a*) Describe in brief bullet points the key elements of your proposed Internet strategy. [5 marks]

(*b*) Draw an outline block diagram of the architecture that would be needed. [5 marks]

(*c*) Make and justify some estimates of the size and cost for its implementation. [5 marks]

(*d*) Outline relevant regulations. [5 marks]

9 Information Retrieval

The PageRank R of a webpage u is defined as:

$$R(u) = (1 - q) + q \sum_{v \in B_u} \frac{R(v)}{N_v}$$

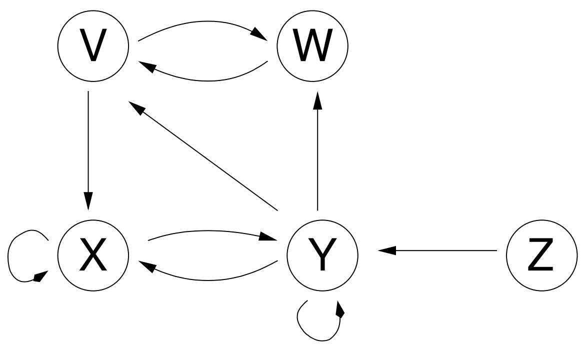
Here, B_u is the set of pages that points to u, N_u is the number of pages that u points to, and q is the probability of staying locally on the webpage.

(*a*) Explain the concept of PageRank, and how it is calculated. [4 marks]

(*b*) Why is it relevant for web search? [3 marks]

(*c*) Give, and briefly explain, the corresponding matrix notation of the PageRank computation. [3 marks]

(*d*) Give the linkage matrix A of the network given in the diagram below. [5 marks]



(*e*) Show the final matrix that will be subjected to the PageRank calculation, if $q = 0.8$ is used. [5 marks]

10 Natural Language Processing

(*a*) Give brief definitions of the following terms. Illustrate the definitions with examples.

(*i*) hyponymy

(*ii*) meronymy

(*iii*) antonymy

[2 marks each]

(*b*) Describe Yarowsky's (1995) technique for word sense disambiguation and illustrate how it would disambiguate the following two senses of "sake":

Sense 1: *sake, interest* (a reason for wanting something done: "for your sake", "died for the sake of his country")

Sense 2: *sake, saki, rice beer* (Japanese alcoholic beverage made from fermented rice, usually served hot)

[14 marks]

11 Security

(*a*) Give *four* uses of anonymous communications other than censorship resistance. [4 marks]

(*b*) Explain the role of latency in anonymous communications. What limits or costs does low latency impose? [4 marks]

(*c*) Imagine you are a government censor, trying to identify which of your citizens are viewing forbidden websites through Tor.

(*i*) If you are able to wiretap the Internet connections of any 1% of the population, what effective capability does this give against Tor users? [4 marks]

(*ii*) If there are currently 1000 active Tor nodes, what extra capability would you acquire if you added a further 100 nodes under your control? Explain any assumptions you make. [4 marks]

(*d*) If you are using Tor to escape censorship, how often should you change the circuit path you use? Explain your answer. [4 marks]

12 Specification and Verification I

(*a*) What is *Russell's Paradox*? Explain how it is avoided in higher-order logic. [4 marks]

(*b*) Explain how the universal quantifier $\forall$ is defined in higher-order logic. Write down a definition of the quantifier and give its type. [4 marks]

(*c*) Define a constant $\mathsf{ExistsFour}$ in higher-order logic such that $\mathsf{ExistsFour}\ P$ represents "$P(x)$ *is true for at least* 4 *values of* x". [4 marks]

(*d*) Define a constant $\mathsf{TotalSpec}$ in higher-order logic that represents the total correctness specification $[P]\ C\ [Q]$. What are the types of the terms representing P, C and Q? [4 marks]

(*e*) Using $\mathsf{TotalSpec}$ from your answer to part (*d*), write down formulae that represent the sequencing and `WHILE` rules for total correctness. [4 marks]

13 System-on-Chip Design

(*a*) How and why does SystemC model registers that have widths not native to the C language? [5 marks]

(*b*) Sketch C++ code in the style of SystemC for a five-bit synchronous counter module that counts up or down depending on an input signal. Explain whether you consider your design to be fully within the nominal "synthesisable subset" of SystemC. [7 marks]

(*c*) Does SystemC reduce the need for RTLs (Register Transfer Languages) such as Verilog and VHDL? Explain what SystemC can provide that is not found in RTLs and *vice versa*. [8 marks]

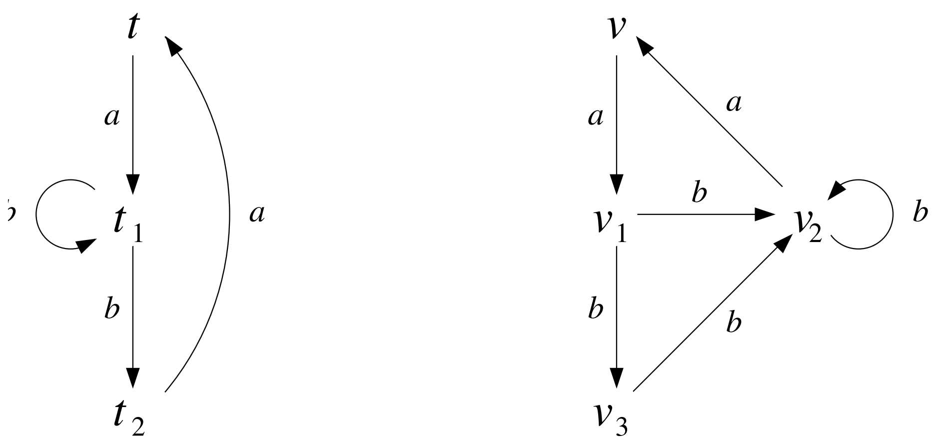
14 Topics in Concurrency

(*a*) In the context of transition systems, explain what a *strong bisimulation* is and what it means for two states to be strongly bisimilar. [3 marks]

(*b*) Describe Hennessy–Milner logic and its semantics. [4 marks]

(*c*) Prove that if two states in transition systems are not strongly bisimilar then there is an assertion of Hennessy–Milner logic that is true of one state but false of the other. [7 marks]

(*d*) In the transition systems below, are states t and v strongly bisimilar? Justify your answer, by exhibiting either a strong bisimulation or an assertion of Hennessy–Milner logic true of one state but false of the other.



[6 marks]

15 Distributed Systems

(*a*) The Domain Name System (DNS) is a large, successful distributed system. Its principal service is to map keys (such as host names) to values.

(*i*) What are the major components of the DNS and how do they interact? [3 marks]

(*ii*) Why does DNS impose a limit on the key length? [2 marks]

(*iii*) Describe *three* aspects of the system that would need to be changed in order to support unlimited key length. [2 marks each]

(*b*) A distributed component-based event middleware supports role-based access control. When the system is started, principal A has role B and principals having role B may carry out action C.

At time T, according to its local clock, a component receives a message that means "principal A no longer has role B." At time $T + 10$ms, this component receives a message that means "perform action C as principal A."

(*i*) Based solely on this information, immediately after the second message is received, why can the component not determine whether A may carry out C? [3 marks]

(*ii*) What data structures should be placed within the two messages described above so that the component can determine, in this situation, whether A may perform C? What algorithms would need to be applied by the middleware to make these data structures useful? [6 marks]

END OF PAPER