COMPUTER SCIENCE TRIPOS Part II

Thursday 5 June 2014 1.30 to 4.30 pm

COMPUTER SCIENCE Paper 9

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
Rough work pad

SPECIAL REQUIREMENTS
Approved calculator permitted

1 Bioinformatics

(*a*) How can you be reasonably confident that an amino acid multiple sequence alignment is optimal? [6 marks]

(*b*) From gene expression analysis we hypothesize that each DNA sequence of a dataset from a different animal species contains at least one similar short subsequence at an unknown position in the dataset. Explain the procedure to identify the short subsequence in the dataset and then cluster the species according to the similarity of the short subsequence. [8 marks]

(*c*) Which algorithms might be useful in aligning very long DNA sequences, such as entire genomes and why? [6 marks]

2 Computer Systems Modelling

(*a*) Consider a Poisson process with rate $\lambda > 0$. Let X_1 be the time of the first event and let X_i be the time between events $(i-1)$-st and i for $i = 2, 3, \ldots$

(*i*) Derive the joint probability distribution of (X_1, X_2). [5 marks]

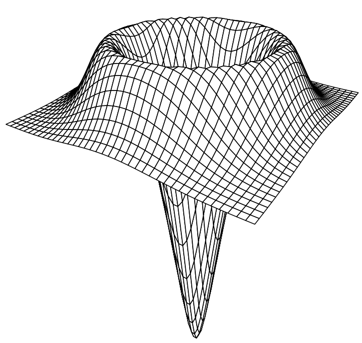
(*ii*) Let $S_n = \sum_{i=1}^{n} X_i$. Derive the probability density function of S_n and give expressions for the mean and variance of S_n. [5 marks]

(*b*) (*i*) Describe what is meant by a FCFS $M/G/1$ queueing system. Your description should include a clear statement of the probabilistic assumptions. [5 marks]

(*ii*) Suppose that you are given a log of timestamps for the arrival and departure events observed in an alleged simulation of a FCFS $M/G/1$ queue with a given general service time distribution. Describe the statistical tests that you would perform on the logged data to test whether the modelling assumptions are satisfied. [5 marks]

3 Computer Vision

(*a*) In early stages of machine vision systems, the isotropic operator shown on the right is often applied to an image $I(x, y)$ in the following way: $[\nabla^2 G_\sigma(x, y)] * I(x, y)$



What is the purpose of this operation? Which class of neurones in the retina does it mimic?

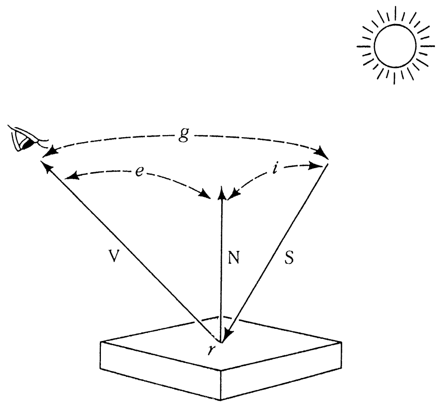
How would the results differ if instead this operation: $G_\sigma(x, y) * \nabla^2 I(x, y)$ were performed; and alternatively if this operation: $\nabla^2 [G_\sigma(x, y) * I(x, y)]$ were performed?

[4 marks]

(*b*) 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 deformable models?

[4 marks]

(*c*) In relation to the image formation diagram shown on the right, explain: (*i*) the concept of a *reflectance map*; (*ii*) what is a *specular* surface; (*iii*) what is a *Lambertian* surface; and (*iv*) what is surface *albedo*. In explaining these concepts, give the defining relationships for the amount of light from a point source that is scattered in different directions by such illuminated surfaces, and describe the inferences that a vision system must make with them.



[4 marks]

(*d*) It can be said that the central problem of pattern recognition is the relative extent of within-class variability and between-class variability. Explain this issue in the context of facial recognition, treating separately the three problems of (*i*) *face detection* (distinguishing faces from non-faces); (*ii*) *face identification;* and (*iii*) *face interpretation* (classifying the expression and pose of the face). How does variability along some of these dimensions influence each of the three problems? Is within-class variability ever helpful, and between-class variability harmful, to the performance of the task?

[8 marks]

4 Denotational Semantics

Let τ be a PCF type.

(*a*) Consider the PCF terms

$$\begin{array}{rcl} \textbf{head} & : & (nat \to \tau) \to \tau \\ \textbf{tail} & : & (nat \to \tau) \to nat \to \tau \\ \textbf{repeat} & : & (nat \to \tau) \to nat \to \tau \end{array}$$

given by the following definitions

$$\begin{array}{rcl} \textbf{head} & = & \textbf{fn}\, s : nat \to \tau.\, s(\mathbf{0}) \\ \textbf{tail} & = & \textbf{fn}\, s : nat \to \tau.\, \textbf{fn}\, n : nat.\, s(\textbf{succ}\, n) \\ \textbf{repeat} & = & \textbf{fix}\Big(\ \textbf{fn}\, f : (nat \to \tau) \to nat \to \tau.\, \textbf{fn}\, s : nat \to \tau.\, \textbf{fn}\, n : nat. \\ & & \qquad \textbf{if}\ (\textbf{zero}\, n)\ \textbf{then}\ (\textbf{head}\, s) \\ & & \qquad \textbf{else if}\ (\textbf{zero}(\textbf{pred}\, n))\textbf{then}\ (\textbf{head}\, s) \\ & & \qquad \textbf{else}\ f\,(\textbf{tail}\, s)\,\big(\textbf{pred}(\textbf{pred}\, n)\big)\ \Big) \end{array}$$

Show that

$$[\![\, \textbf{fn}\, s : nat \to \tau.\, \textbf{tail}\big(\textbf{tail}(\textbf{repeat}\, s)\big)\,]\!] \;=\; [\![\, \textbf{fn}\, s : nat \to \tau.\, \textbf{repeat}(\textbf{tail}\, s)\,]\!]$$

in the domain $\big((\mathbb{N}_\perp \to [\![\tau]\!]) \to (\mathbb{N}_\perp \to [\![\tau]\!])\big)$. [6 marks]

(*b*) Define a closed PCF term

$$\textbf{shuffle} : (nat \to \tau) \to (nat \to \tau) \to nat \to \tau$$

such that

$$\begin{array}{rcl} [\![\textbf{head}]\!]\big([\![\textbf{shuffle}]\!]\ s\ t\big) & = & [\![\textbf{head}\ s]\!] \\ [\![\textbf{tail}]\!]\big([\![\textbf{shuffle}]\!]\ s\ t\big) & = & [\![\textbf{shuffle}]\!]\ t\ ([\![\textbf{tail}]\!]\ s) \end{array}$$

for all $s, t \in (\,\mathbb{N}_\perp \to [\![\tau]\!]\,)$. Briefly justify your answer. [5 marks]

(*c*) (*i*) Define the notion of least pre-fixed point $\mathit{fix}(f)$ in a domain D of a continuous function f in the function domain $(D \to D)$. [3 marks]

(*ii*) Prove that

$$[\![\textbf{repeat}]\!] \sqsubseteq [\![\textbf{fn}\, s : nat \to \tau.\, \textbf{shuffle}\, s\, s]\!]$$

in the domain $\big((\mathbb{N}_\perp \to [\![\tau]\!]) \to (\mathbb{N}_\perp \to [\![\tau]\!])\big)$. [6 marks]

5 Digital Signal Processing

A discrete sequence $\{x_n\}$ can be converted into a continuous representation

$$\hat{x}(t) = t_{\mathrm{s}} \cdot \sum_{n=-\infty}^{\infty} \delta(t - n \cdot t_{\mathrm{s}}) \cdot x_n,$$

where t_{s} is the sampling period.

(*a*) State two characteristic properties of Dirac's δ function. [2 marks]

(*b*) Describe briefly how this representation helps to explain aliasing. [4 marks]

(*c*) Define three functions $h(t)$, such that convolving $\hat{x}(t)$ with $h(t)$ results in

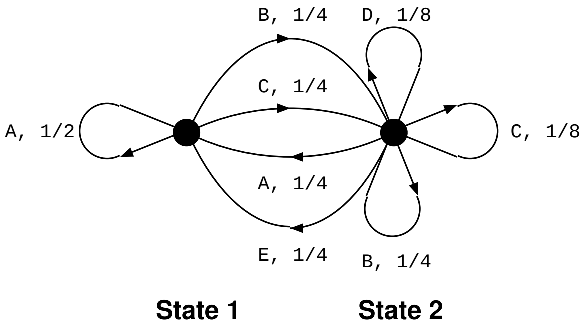
(*i*) the output of an idealized analog-to-digital converter that holds the output voltage of each sample x_n for the time interval from $t = n \cdot t_{\mathrm{s}}$ until the next sample x_{n+1} arrives at time $t = (n+1) \cdot t_{\mathrm{s}}$; [4 marks]

(*ii*) linear interpolation of $\{x_n\}$; [4 marks]

(*iii*) reconstruction of a signal $x(t)$ that was sampled as $x_n = x(n \cdot t_{\mathrm{s}})$, assuming that the Fourier transform of $x(t)$ is zero at any frequency f with $|f|^{-1} \le t_{\mathrm{s}}$ or $|f|^{-1} \ge 2t_{\mathrm{s}}$. [6 marks]

6 Information Theory and Coding

(*a*) A two state Markov process emits the letters {A, B, C, D, E} with the probabilities shown for each state. Changes of state can occur when some of the symbols are generated, as indicated by the arrows.



(*i*) What are the state occupancy probabilities? [1 mark]

(*ii*) What is the probability of the letter string AD being emitted? [1 mark]

(*iii*) What is the entropy of State 1, what is the entropy of State 2, and what is the overall entropy of this symbol generating process? [5 marks]

(*b*) A fair coin is secretly flipped until the first head occurs. Let X denote the number of flips required. The flipper will truthfully answer any "yes-no" questions about his experiment, and we wish to discover thereby the value of X as efficiently as possible.

(*i*) What is the most efficient possible sequence of such questions? Justify your answer. [2 marks]

(*ii*) On average, how many questions should we need to ask? Justify your answer. [2 marks]

(*iii*) Relate the sequence of questions to the bits in a uniquely decodable prefix code for X. [1 mark]

(*c*) Define complex Gabor wavelets, restricting yourself to one-dimensional functions if you wish, and list four key properties that make such wavelets useful for encoding and compressing information, as well as for pattern recognition. Explain how their self-Fourier property and their closure under multiplication (i.e. the product of any two of them is yet again a Gabor wavelet) gives them also closure under convolution. Mention one disadvantage of such wavelets for reconstructing data from their projection coefficients. [8 marks]

7 Mobile and Sensor Systems

(*a*) (*i*) Compare the infrastructure-based and infrastructure-less modalities of the Wi-Fi (802.11) standard. [2 marks]

(*ii*) Design an efficient infrastructure-less system that displays the slides of a lecture to the tablets (such as iPads, etc.) of the students. The slides are initially stored only on the lecturer's laptop. Describe briefly the architecture of the system, focussing on the communication aspects. [4 marks]

(*iii*) Introduce and describe an efficient mechanism for allowing students to add annotations and comments to the slides during the lecture. The annotations should also propagate on all students' copies. [4 marks]

(*b*) (*i*) Define the principles of "participatory sensing" and describe an example of it. [2 marks]

(*ii*) Discuss an application using the smartphone's microphone as a sensor and the issues related to performance and device constraints. [4 marks]

(*iii*) Assume that you have to build an application that periodically collects data from the phone's accelerometer sensor and uses this information to classify user activities. The result of the classification is then sent to a back-end server. You should assume that the classifiers run on the phone. Describe the strategies you would adopt in order to reduce the energy consumption. [4 marks]

8 Natural Language Processing

(*a*) Use the following text (from Eliot's 'The Hollow Men') to derive distributions for *eyes*, *here* and *valley*. Use a 5-word window including open- and closed-class words, ignore case and line-breaks and weight contexts by frequency.

> The eyes are not here
> There are no eyes here
> In this valley of dying stars
> In this hollow valley
> This broken jaw of our lost kingdoms

[5 marks]

(*b*) What is cosine similarity? Show how to calculate the cosine similarity between each pair of the distributions you produced above (i.e., eyes/here, eyes/valley, valley/here). [5 marks]

(*c*) The following table shows some similarities for pairs of nouns calculated using distributions extracted from the British National Corpus.

ring	bracelet	0.64	ring	cat	0.13
ring	necklace	0.55	lord	knight	0.37
ring	finger	0.54	lord	ring	0.41

Discuss what these results suggest about distributional similarity with respect to human similarity judgements and with respect to the concept of synonymy used in lexical semantics. [4 marks]

(*d*) Describe the notion of textual entailment, as investigated in the Recognising Textual Entailment (RTE) challenge. Outline how one could investigate the use of distributional methods to improve the performance of a bag-of-words method for this task. [6 marks]

9 Optimising Compilers

(*a*) Explain the idea of an effect system (also known as a type and effect system), including a typical judgement form $\Gamma \vdash e : t, F$ explaining how this differs from a traditional ML-like type system. Use the words 'immediate' and 'latent' in your answer. [3 marks]

(*b*) Give (*i*) the set of effects and (*ii*) the inference rules for an effect system for a language based on lambda-calculus but including constants, if-then-else, let-in, and reads and writes to channels ξ. The language types should consist of integers and functions. At this point you should only give effect rules corresponding to standard typing rules for your language (i.e. one per syntactic form). [5 marks]

(*c*) Now give an expression e which is well-typed without effects, but which fails to have an effect judgement of the form $\Gamma \vdash e : t, F$. Explain how an additional rule, similar to subtyping, enables this program to be given an effect judgement. [4 marks]

(*d*) Suppose now your language contains an infix constant '`+`' representing curried addition with left-to-right evaluation of its operands. We would like to be able to optimise so as to execute e_1 and e_2 in parallel *when this is safe* before adding their results. Give a (safe but not unreasonably restrictive) criterion on the effects of e_1 and e_2 for this optimisation. Discuss any issues that arise in the case that e_1 and e_2 share a common channel. [4 marks]

(*e*) Finally consider, *using your rules from* Parts (*b*) *and* (*c*), whether your criterion from Part (*d*) can successfully (*i*) ascribe a type and effect to, and (*ii*) parallelise, the two calls to `apply` in the program:

```
let apply = λf. λx. f x
in let add1 = λx. x+1
in let add2 = λx. (print!42. x+2)
in apply add1 10 + apply add2 20
```

where `print!`e_1`.`e_2 evaluates e_1 printing its value to standard output, then evaluates e_2 yielding its value as result.

Either justify your parallelisation or suggest a means of avoiding any issue inhibiting parallelisation. [4 marks]

10 Principles of Communications

(*a*) The Internet provides a "best effort" service in forwarding packets, by default. Explain what this means in terms of performance of the system, making due reference to statistical multiplexing of resources, and to the end-to-end fairness properties (or lack of them) that may result. [10 marks]

(*b*) Many communications systems are designed as a set of layered services. Discuss the advantages and disadvantages of layering in communications systems in terms of the trade-offs between flexibility and performance that this approach may provide. Illustrate your answer with examples. [10 marks]

11 System-on-Chip Design

(*a*) A mobile phone uses an average power of 20 mW when not in use. Given that its battery voltage is 3V with capacity 500 mAh, what is the maximum interval between rechargings? [2 marks]

(*b*) The supply regulator converts the battery voltage to a 1 Volt supply for the digital logic and is 80% efficient. If the phone dissipation during a CPU-bound computer game is 3 Watts and the processor clock frequency is 500 MHz, what, roughly, is the average track capacitance discharged each clock cycle? State any assumptions. For instance, you might assume that the screen backlight dissipates 1 Watt. [3 marks]

(*c*) What are the main differences, in terms of latency, throughput, capacity and energy use, between an on-chip register file and an off-chip RAM? [4 marks]

(*d*) Simulation in advance of manufacture can provide an estimate of power consumption. Explain what power estimation procedures might be used when simulating at each of the following design stages:

(*i*) A TLM model of the SoC that is running the production software.

(*ii*) A Verilog RTL model of the complete SoC prior to gate synthesis.

(*iii*) A Verilog netlist of the complete SoC prior to placement.

(*iv*) A Verilog netlist of the complete SoC after place and route.

[8 marks]

(*e*) A design goal for the phone is, "The battery life on standby shall always be greater than three days." Discuss whether this is a safety or liveness assertion and where it might best be embodied in the design flow. [3 marks]

12 Topical Issues

(a) Give five differences between a passive UHF RFID tag and a standard printed barcode. [5 marks]

(b) Consider a group of four 4-bit passive UHF RFID tags with the identifiers {1011, 1010, 0111, 1001}.

(i) Give the sequence of queries necessary to enumerate the group of four tags when using the Binary Tree Walking Algorithm (BTWA). [3 marks]

(ii) Show how the Query Tree Walking Algorithm (QTWA) can fully enumerate the same four tags using just four queries. [4 marks]

(iii) Discuss the extent to which a passive eavesdropper can identify the RFID tags in range of a given reader for both BTWA and QTWA. Consider eavesdroppers that are very close to the reader and also eavesdroppers near the maximum transmission range of the reader. Illustrate your answer using your solutions to (b)(i) and (b)(ii). [8 marks]

13 Types

(*a*) Give an account of the Polymorphic Lambda Calculus (PLC). You should define the PLC types, expressions and typing environments, give the PLC typing rules, and define the relations of *beta-reduction* and *beta-conversion*. Explain why beta-conversion is a decidable relation for expressions that are typeable. [10 marks]

(*b*) Find, with justification, a PLC type τ for which the following typings are both provable:

$$\{\,\} \vdash \Lambda\,\alpha, \beta\,(\lambda x : \alpha(\lambda y : \beta(\Lambda\,\gamma\,(\lambda z : \tau(z\,x\,y))))) : \forall\,\alpha, \beta\,(\alpha \to (\beta \to \forall\,\gamma\,(\tau \to \gamma)))$$

$$\{\,\} \vdash \Lambda\,\alpha, \beta\,(\lambda z : \forall\,\gamma\,(\tau \to \gamma)(z\,\alpha\,(\lambda x : \alpha(\lambda y : \beta(x)))))$$
$$: \forall\,\alpha, \beta\,((\forall\,\gamma\,(\tau \to \gamma)) \to \alpha)$$

[5 marks]

(*c*) Give infinitely many different closed PLC expressions in beta-normal form of type $\forall\,\alpha\,((\alpha \to \alpha) \to (\alpha \to \alpha))$. [2 marks]

(*d*) Use your answer to part (*a*) to show that there is no closed PLC expression Y of type $\forall\,\alpha\,((\alpha \to \alpha) \to \alpha)$ for which the beta-reduction $Y\,\alpha\,f \to f(Y\,\alpha\,f)$ holds. [3 marks]

END OF PAPER