UNIVERSITY OF CAMBRIDGE

COMPUTER SCIENCE TRIPOS Part II

Wednesday 1 June 2016 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.*

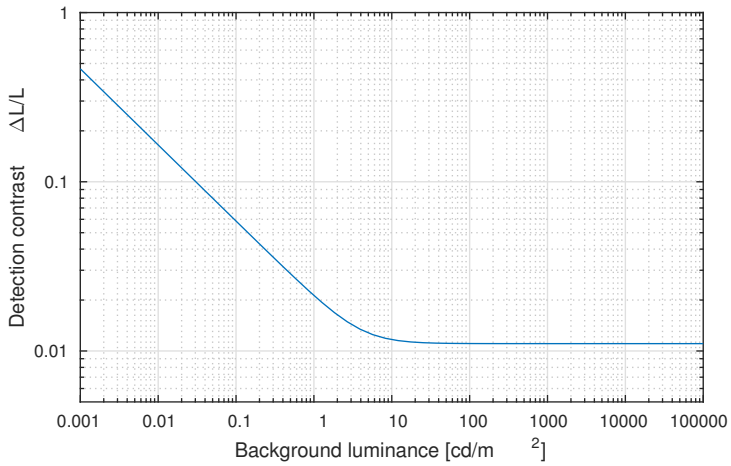
You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

STATIONERY REQUIREMENTS
Script paper
Blue cover sheets
Tags

SPECIAL REQUIREMENTS
Approved calculator permitted

1 Advanced Graphics

(a) Answer the questions below, referring to the contrast-versus-intensity function in the plot shown below when appropriate.



(i) Why is luminance usually plotted using the logarithmic scale? [2 marks]

(ii) Why are stars visible at night but not in the day-time? Justify your answer referring to the contrast-versus-intensity plot above. [3 marks]

(iii) Is the sensitivity of the visual system higher when it operates in a bright environment or in a dim environment? Explain why. [3 marks]

(iv) Why is the power of 1/3 function used in CIE Lab and CIE Luv colour spaces? [2 marks]

(b) You are given a gray-scale high dynamic range image I represented in absolute luminance units of cd/m^2. The image is to be viewed on a target display with the peak luminance $L_{\text{peak}} = 100\,\text{cd/m}^2$ and the black level $L_{\text{black}} = 1\,\text{cd/m}^2$ (both values measured in the dark environment). The display is viewed in a bright environment and the amount of the light reflected from the screen was measured to be $L_{\text{refl}} = 2\,\text{cd/m}^2$.

(i) Calculate the effective dynamic range (contrast) of that display in the bright environment and express it as a contrast ratio N:1. Show the formula as well as the final answer. [2 marks]

(*ii*) What operation needs to be performed on the image I to make it twice as bright? Express the final result as luminance. [2 marks]

(*iii*) How can the contrast of image I be reduced by a factor of 2 so that the luminance values equal to L_{peak} do not change? Express the final result as luminance. [3 marks]

(*iv*) Write pseudo-code that adds glare to image I. The glare is modelled as a point spread function g. Your formula must exclude the glare that is naturally produced in the eye when viewing the target display described above. Use the $*$ symbol for the convolution operator. Express the final result as luminance. [3 marks]

2 Artificial Intelligence II

You have been given a set $\mathbf{s}$ containing m labelled training examples for a binary classification problem. The ith example is $(\mathbf{x}_i, y_i)$. Approximately 10% of the training examples are labelled $+1$ and the remainder -1. You wish to use $\mathbf{s}$ to train a classifier $h(\mathbf{x}; \mathbf{w}, p)$, where $\mathbf{w}$ is a set of weights and p is a hyperparameter. Your colleague advises you to ignore p and simply set it to 1, and then to use some algorithm to choose $\mathbf{w}$ to minimize the quantity

$$E(\mathbf{w}) = \sum_{i=1}^{m} I(h(\mathbf{x}_i; \mathbf{w}, 1) \neq y_i)$$

where I is the indicator function $I(x) = 1$ if x is true and $I(x) = 0$ otherwise. Your colleague also suggests that the value of $E(\mathbf{w})$ that results should be used as an assessment of your classifier's performance.

(a) Identify *three* errors in your colleague's advice, in each case explaining why it is erroneous. [6 marks]

(b) Suggest a more appropriate way of training and assessing the classifier. In your answer you should address each of the three reasons given in Part (a) and provide an alternative approach, defining in full any new concepts that you introduce. [6 marks]

(c) In addition, identify one further way in which your colleague's suggested procedure is deficient. How would you correct this? [2 marks]

(d) A second colleague points out that rather than choosing a specific $\mathbf{w}$ you should consider a fully Bayesian approach, taking all possible values for $\mathbf{w}$ into account by computing $\Pr(y = +1|\mathbf{x})$. Give *two* reasons why this approach might be preferred, and *two* reasons why it might not, explaining your answer in each case. [6 marks]

3 Comparative Architectures

(*a*) Describe briefly six factors which might influence or constrain the design of a new processor. [6 marks]

(*b*) The performance of a superscalar processor is often enhanced with hardware to support the following:

- branch prediction
- register renaming
- out-of-order execution
- the speculative reordering of load instructions
- strided prefetching

(*i*) Sketch an assembly language program that would benefit from the use of all of these techniques when executed on a superscalar processor. Briefly describe how each of the techniques helps to improve the performance of your program. [10 marks]

(*ii*) Briefly outline two example programs for which the adoption of the techniques listed would not provide a significant performance improvement. [4 marks]

4 Computer Systems Modelling

(*a*) (*i*) Suppose that $F_X(x)$ is a distribution function. Show the *inverse transform result*, namely that, if U is a random variable uniformly distributed in the interval $(0, 1)$ then

$$X = F_X^{-1}(U)$$

is a random variable with distribution function $\mathbb{P}(X \leq x) = F_X(x)$. [4 marks]

(*ii*) Discuss the notion of a pseudo-random number generator for uniform random variables. Describe suitable algorithms for generating pseudo-random numbers. [6 marks]

(*iii*) Using the inverse transform result in part (*a*)(*i*) derive a method to generate a stream of independent pseudo-random numbers from an exponential distribution with parameter $\lambda > 0$. What are the true mean and variance of these numbers in terms of λ? [4 marks]

(*b*) (*i*) Suppose that you conduct a simulation experiment to estimate the mean, μ, of a random quantity X from a sample of n values $X_1, X_2, \ldots, X_n$. How would you estimate μ? [2 marks]

(*ii*) Now suppose that your simulation also yields a sample of n values $Y_1, Y_2, \ldots, Y_n$ of the random quantity Y where $\mathbb{E}(Y) = \mu_Y$ is a known number. How would you use the method of *control variates* to improve your estimator of μ? Your answer should mention all quantities that may need to be estimated and in what way you will improve the estimation of μ. [4 marks]

5 Computer Vision

(*a*) Explain the key elements in "FaceNet" that enabled it to achieve a major breakthrough in face recognition performance, with impressive pose-invariance as illustrated below, and illumination invariance. [5 marks]

(*b*) Explain the Bayer pattern of colour separation used in many sensors for image acquisition, with reference to its effects on spatial resolution, the relative importance of luminance versus chrominance resolution, and the relative pixel densities of the various colour planes. [3 marks]

(*c*) Consider the following pair of (6×6) filter kernels:

-1	-1	2	2	-1	-1
-1	-3	4	4	-3	-1
-1	-4	5	5	-4	-1
-1	-4	5	5	-4	-1
-1	-3	4	4	-3	-1
-1	-1	2	2	-1	-1

1	-1	-1	1	1	-1
1	-2	-3	3	2	-1
1	-3	-4	4	3	-1
1	-3	-4	4	3	-1
1	-2	-3	3	2	-1
1	-1	-1	1	1	-1

(*i*) Why do these two kernels constitute a quadrature pair? [2 marks]

(*ii*) To what kinds of image structure, and which orientations, are these detector kernels most sensitive? [2 marks]

(*iii*) How would these kernels be applied directly to an image for filtering or feature extraction? [2 marks]

(*iv*) How could their respective Fourier Transforms alternatively be applied to an image, to achieve the same effect as in (*iii*) but faster? [2 marks]

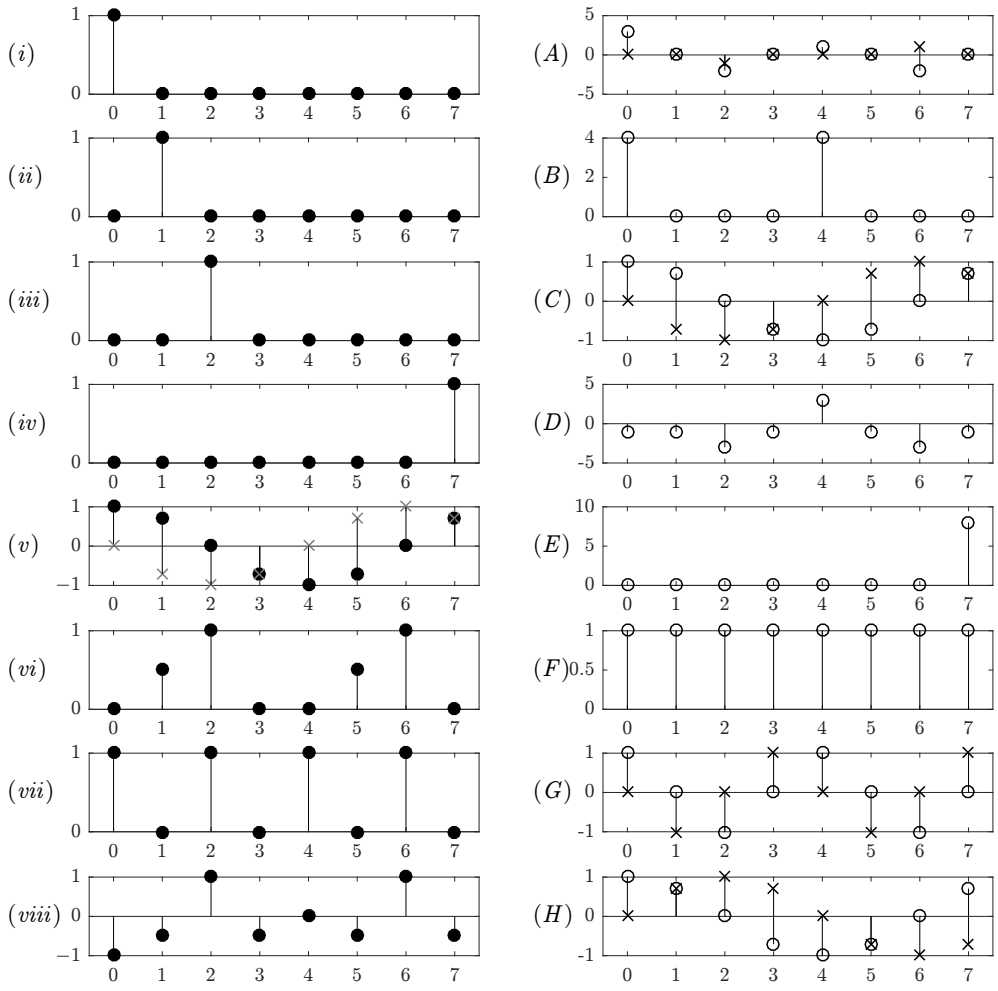
(*v*) What is the "DC" response of each of the kernels, and what is the significance of this? [2 marks]

(*vi*) How could these kernels be combined to locate facial features? [2 marks]

6 Digital Signal Processing

(*a*) Figures (*i*)–(*viii*) show eight different input vectors $x \in \mathbb{C}^8$. For each, identify one of figures (*A*)–(*H*) that shows the DFT output $X \in \mathbb{C}^8$ with $X_k = \sum_{n=0}^{7} x_n \cdot \mathrm{e}^{-2\pi \mathrm{j} kn/8}$.

Briefly explain each choice. Real components are shown as circles. For non-real vectors, the imaginary components are shown in addition as crosses. [8 marks]



(*b*) Are these statements true or false? Explain your answers. [3 marks each]

(*i*) The system $y_n = x_n + y_{n-1}$ has an impulse response with z-transform $\frac{1}{1+z}$.

(*ii*) A continuous signal can *only* be reconstructed after sampling if the sampling frequency is larger than twice the highest frequency in the signal.

(*iii*) Convolution of a signal with a triangular window function causes its power spectrum to be multiplied with a sinc^3 function.

(*iv*) To convert the z-transform $H(z)$ of the impulse response of any LTI filter into the z-transform of its step response, divide $H(z)$ by $1 - z^{-1}$.

7 E-Commerce

(*a*) Define *fungibility* in the context of an online game currency. [5 marks]

(*b*) Discuss the advantages and disadvantages of making an online game currency fungible. [5 marks]

(*c*) Discuss the management of the game's internal economy. [5 marks]

(*d*) What are some advantages or disadvantages of using block chain technology to implement online game currencies? [5 marks]

8 Information Retrieval

Consider the following documents:

doc_1	phone ring person happy person
doc_2	dog pet happy run jump
doc_3	cat purr pet person happy
doc_4	life smile run happy
doc_5	life laugh walk run run

(*a*) (*i*) Construct the inverted index required for ranked retrieval for these five documents. Assume that no stemming or stop-word removal is required. [3 marks]

(*ii*) What is the complexity of processing a two-term conjunctive query using standard postings lists? Briefly describe one technique that can improve this efficiency. [2 marks]

(*iii*) Relating to the sample documents above, outline how the processing of the following Boolean query can be optimised:

happy AND run AND pet [2 marks]

(*iv*) What is the query-likelihood method in the language modelling approach to information retrieval? How does this differ conceptually from the measure of similarity used in the vector space model? [3 marks]

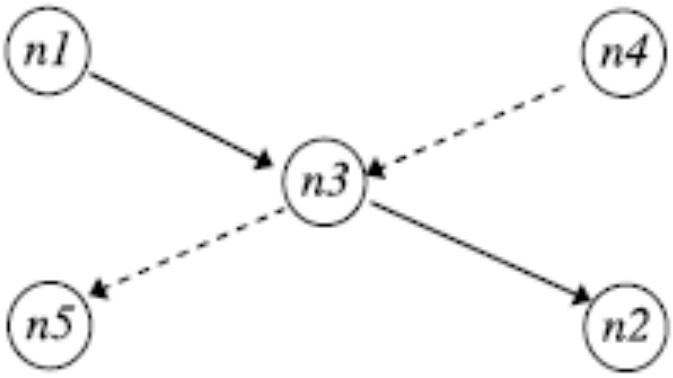
(*b*) (*i*) Smoothing is crucial in the language modelling approach to information retrieval. Why is smoothing important and how is it typically achieved? [2 marks]

(*ii*) Given the query {*happy person smile*}, show how a unigram language modelling approach would rank the documents outlined above. Choose a suitable form of smoothing and include all your workings. State any other assumptions made. [6 marks]

(*iii*) How might you relax the *term-independence* assumption in the unigram language model and how might it affect subsequent retrieval? [2 marks]

9 Principles of Communications

(*a*) Network Coding can be added to the routing and forwarding layer in a wireless mesh network to improve the capacity. The gains come from reduced transmissions, and from the associated reduction in contention for the radio channel. Show with examples how the coding scheme that combines packets in the simplest topology illustrated below can increase capacity by 1/3. If we increase the number of flows traversing the central node, how does the capacity improve? In the limit, for a star network, how good can the improvement be?



[15 marks]

(*b*) The standard version of the Transmission Control Protocol recovers from lost packets by using positive acknowledgements coupled with timeout or duplicate acknowledgements to trigger retransmissions. Describe briefly how network coding can be applied across multiple data packets instead, explaining how the acknowledgement scheme must be modifed to accommodate coded packets.

[5 marks]

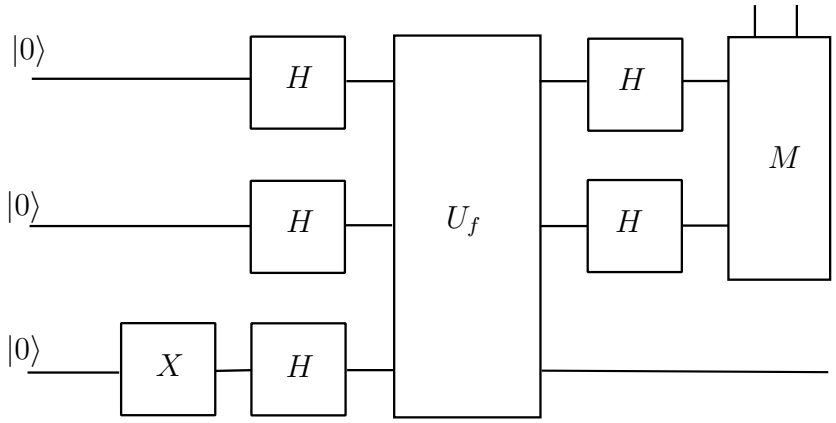
10 Quantum Computing

Let $f : \{0,1\}^2 \to \{0,1\}$ be a Boolean function of two inputs. Let U_f be the implementation of f as a unitary operator on 3 qubits defined by:

$$U_f|x\rangle|y\rangle|z\rangle = |x\rangle|y\rangle|z \oplus f(x,y)\rangle,$$

where $\oplus$ denotes the exclusive-or operation, and $|x\rangle|y\rangle|z\rangle$ is any computational basis state.

Consider the following circuit (a two-qubit version of the Deutsch-Josza circuit) in which X denotes a NOT gate, H denotes a Hadamard gate and M is a two-qubit measurement in the computational basis.



(*a*) Show that if f is a constant function, the outcome of the measurement M is 00 with probability 1. [6 marks]

(*b*) Show that if f is the XOR function, the outcome of the measurement M is 11 with probability 1. [6 marks]

(*c*) What are the probabilities of M measuring 00 and 11 respectively, if f is the Boolean AND function? [8 marks]

11 Security II

(*a*) Why does the formal security definition for collision-resistant hash functions require a key s and a security parameter n, even though most commonly used standard secure hash functions lack such input parameters? [4 marks]

(*b*) If $h_s : \{0,1\}^* \to \{0,1\}^{\ell(n)}$ is a collision-resistant hash function, do the following constructions H_s also provide collision-resistant hash functions? Explain your answers. [2 marks each]

(*i*) $H_s(x) = h_s(x) \parallel x$ (i.e. append x)

(*ii*) $H_s(x) = h_s(x) \parallel \mathrm{LSB}(x)$ (i.e. append least significant bit of x)

(*iii*) $H_s(x) = h_s(x \mid 1)$ (bitwise-or, i.e. set least significant bit of x to 1)

(*c*) Use Euler's theorem to calculate $5^{-1} \bmod 8$. [4 marks]

(*d*) The standard Digital Signature Algorithm (DSA) uses a cyclic subgroup $\mathbb{G} \subset \mathbb{Z}_p^*$ of the integers modulo a prime p, with prime order q, where q divides $p-1$.

(*i*) Give two advantages of using a multiplicative subgroup of prime order, as opposed to just using $\mathbb{Z}_p^*$, in cryptographic schemes based on the Discrete Logarithm problem. [2 marks]

(*ii*) Why is it possible to choose q substantially smaller than p, and what is an advantage of doing so? [4 marks]

12 System-on-Chip Design

(*a*) SoC design involves dividing work between hardware and software as well as deciding the number of general-purpose and custom processors and co-processors to be used. What main factors influence these design decisions and the associated manual partition of envisaged workload over these resources? [5 marks]

(*b*) A lossless compression algorithm converts fixed-size, 1 kByte blocks of data to a variable-sized block that is normally shorter. Suggest a suitable signature (argument and return types) for a software function that implements the algorithm. Draw a diagram showing the external wiring to the neighbouring SoC components for an appropriate, high-performance, hardware implementation. State any assumptions that guide your design approach. [6 marks]

(*c*) A synchronous hardware module has separate input and output ports that each convey 32-bit words. Handshaking is required on both ports since it is unpredictable whether the module or its environment are ready to exchange data in either direction at any time. Give a diagram or RTL module definition for such a component and precisely describe the handshaking protocol in words. State the maximum throughput of your protocol. [5 marks]

(*d*) Why might it be useful to have a formal specification of your protocol from part (*c*) during design and testing? What, if anything, might we infer about the number of words stored inside the module from the protocol? [4 marks]

13 Topical Issues

(*a*) Compare and contrast the Internet of Things (IoT) with the conventional internet. [4 marks]

(*b*) Describe how Bluetooth Low Energy (BLE) was designed to address the following IoT-related issues. Your answers should consider both BLE advertising and data channels.

(*i*) Reduce the overall power consumption of peripherals. [6 marks]

(*ii*) Handle radio channel contention from other IoT devices. [6 marks]

(*iii*) Handle radio channel interference from other users of the same 2.4 GHz radio band such as WiFi. [4 marks]

14 Topics in Concurrency

(a) Define the *token game* for basic Petri nets. [3 marks]

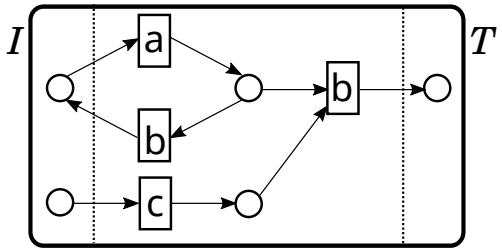
(b) When is a basic Petri net *safe* from an initial marking? [2 marks]

(c) An LB-net is a basic Petri net $(B, E, pre, post)$ accompanied by

- a labelling function $\lambda : E \to Act$ from its events to a set of actions Act
- subsets of conditions $I \subseteq B$ and $T \subseteq B$. The initial conditions I are marked when the process starts and the terminal conditions T are marked when the process has terminated.

LB-nets are drawn with labels inside events and boxes surrounding the initial and terminal conditions.

(i) Draw the labelled transition system of the following LB-net. The initial state should correspond to the initial conditions I being marked and labels on the transitions should correspond to actions, not events. [4 marks]



(ii) Ignoring the particular sets that states represent, is there an LB-net with an *injective* labelling function λ that gives rise to the same labelled transition system? Justify your answer briefly. [2 marks]

(iii) A simple process language has the following syntax.

$$p ::= \alpha \mid p + p' \mid p \parallel p' \mid p; p'$$

where $\alpha \in Act$. As in CCS, $+$ represents the nondeterministic sum of processes and $\parallel$ represents the parallel composition. The process $p; p'$ represents the sequential composition of p and p.

Draw diagrams to describe the inductive definition of an LB-net semantics for this fragment. [7 marks]

(iv) An iteration operator p^* is proposed with LB-net semantics such that its sets of initial and terminal conditions are equal: $I = T$. Discuss briefly how this affects the semantics you gave in part (c)(iii). [2 marks]

15 Types

(*a*) In Mini-ML, define the relation of *specialisation* between

(*i*) type schemes and types, $\forall A\,(\tau) \succ \tau'$ [1 mark]

(*ii*) type schemes and type schemes, $\forall A\,(\tau) \succ \forall A'\,(\tau')$ [2 marks]

(*b*) What is meant by the *principal* type scheme of a closed expression in Mini-ML? [2 marks]

(*c*) State the *Hindley-Damas-Milner Theorem* for the Mini-ML typeability problem. [2 marks]

(*d*) Define what is meant by a Mini-ML *typing problem*. Outline a type inference algorithm for Mini-ML that operates on typing problems. You should explain what is a *solution* and a *principal solution* for a typing problem, state the properties of the output of the algorithm and explain its overall structure. How does the algorithm make use of fresh type variables and of unification? Illustrate your answer by describing how the algorithm operates on function application expressions. [13 marks]

END OF PAPER