COMPUTER SCIENCE TRIPOS Part II

Tuesday 2 June 2020 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 Algorithms

(a) State the fundamental theorem of Linear Programming. [3 marks]

(b) Consider the following linear program:

$$\begin{aligned} \text{minimise} \quad & 4 \cdot x_1 - x_2 \\ & -x_1 + 5x_2 \geq 4 \\ & x_1 - 0.5x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

(i) Convert this linear program into slack form. [3 marks]

(ii) What is the number of different slack forms of the linear program in Part (b)(i)? [2 marks]

(iii) Give at least one non-feasible and one feasible basic solution of the linear program in (b)(i). [4 marks]

(c) Consider the following separation problem. We are given m points $x^1 = (x_1^1, x_2^1), x^2 = (x_1^2, x_2^2), \ldots, x^m = (x_1^m, x_2^m) \in \mathbb{R}^2$ and n points $y^1 = (y_1^1, y_2^1), y^2 = (y_1^2, y_2^2), \ldots, y^n = (y_1^n, y_2^n) \in \mathbb{R}^2$. The goal is to find a "separating" vector $w = (w_1, w_2) \in \mathbb{R}^2$ (if it exists) such that:

$$\langle x^i, w \rangle = \sum_{j=1}^{2} x_j^i w_j > 0 \qquad \text{for } i = 1, 2, ..., m,$$

and

$$\langle y^i, w \rangle = \sum_{j=1}^{2} y_j^i w_j < 0 \qquad \text{for } i = 1, 2, ..., n.$$

(i) Create a new, equivalent system of inequalities by replacing each strict inequality by a suitable non-strict inequality. Justify why this new system has a solution if and only if the original system has one. [4 marks]

(ii) Based on your answer in Part (c)(i), how can you solve the above problem using linear programming? [4 marks]

2 Bioinformatics

(*a*) Describe the time and space complexity for finding the global alignment between two DNA sequences when they are very similar. [4 marks]

(*b*) Describe how match, mismatch and gap penalty (initiation and elongation) affect the score in global sequence alignment. [4 marks]

(*c*) Describe, with one example, how the number of mutations affects the phylogenetic analysis using the distance algorithm. [4 marks]

(*d*) Describe the differences in algorithmic complexity between the distance and parsimony phylogenetic methods. [4 marks]

(*e*) Describe why the Viterbi algorithm can help identify protein structural parts (alpha helix, beta sheet, coil) in a sequence of amino acids and describe how it works. [4 marks]

3 Comparative Architectures

(*a*) (*i*) For each of the processors described below outline a possible microarchitecture. Include a labelled block diagram that illustrates the main components of the core pipeline and how they are interconnected. Individual pipeline stages should be shown with a brief description of their contents.

(A) A simple superscalar processor with a short pipeline. It can fetch and issue up to two instructions per cycle and instructions are issued in program order. [4 marks]

(B) A high-performance superscalar processor that supports out-of-order execution. It is able to fetch and issue up to 6 instructions per cycle. It has a deep pipeline and aims to support a high clock frequency. [6 marks]

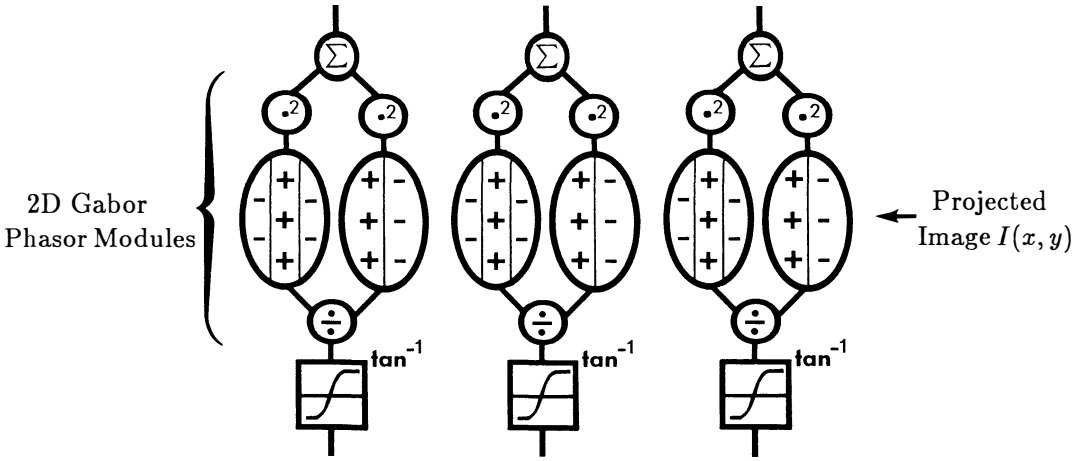
(*ii*) In practice, the area of these types of processor may differ by a factor of five or more. What contributes to this large difference in area? [5 marks]

(*b*) An indirect branch may have multiple target addresses associated with it. Why is this problematic for a simple Branch Target Buffer (BTB) design? [2 marks]

(*c*) If indirect branches favour a particular branch target, and only infrequently branch to other targets, how might the design of the BTB be optimised? [3 marks]

4 Computer Vision

(a) Neural receptive fields used in early stages of vision can be regarded as linear integro-differential operators of the first- and second-orders, represented by the elongated ovals within the diagram below. Explain how they can be used for oriented edge detection, and also state the basis for a Fourier interpretation of them as anisotropic bandpass filters. Explain how combining their outputs by the nonlinear operations depicted in the rest of this diagram (sum-of-squares, and response ratio) can be used for higher-level feature detection. [8 marks]

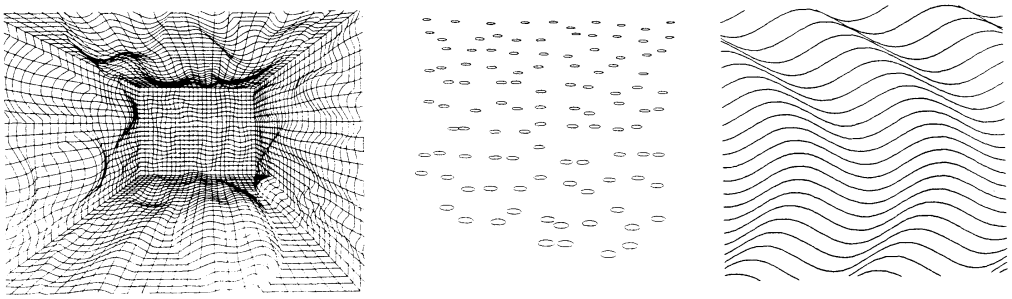


(b) In self-driving cars, the following acronyms are names for automated vision systems. Define them and briefly describe how they work.

(i) LIDAR [3 marks]

(ii) SLAM [3 marks]

(c) Discuss the use of texture gradients as a depth cue in computer vision. How can texture gradients be measured? What prior assumptions are needed to make computations about depth and shape possible? You may find it helpful to refer to the following texture examples. [6 marks]



5 Cryptography

(*a*) (*i*) One way to use a secure hash function H to form a message-authentication code is the construct $\mathsf{Mac}_K(M) = H(K\|M)$. What problem with that approach does the HMAC construct solve? [4 marks]

(*ii*) Why does the HMAC construct pad the key? [2 marks]

(*b*) Your opponent has started using *HomeBrew*, a new block cipher $C = E_K(M)$ that they invented last week. It uses a 96-bit key $K = K_1\| \ldots \|K_{12}$, where each of the 12 bytes K_i ($1 \leq i \leq 12$) is used as an 8-bit subkey in one of the 12 rounds that apply a keyed permutation f:

$R_0 := M$
for $i := 1$ **to** 12
  $R_i := f_{K_i}(R_{i-1})$
$C := R_{12}$

Describe an attack to find K for this type of block cipher that is practical for an adversary with a computer fast enough to execute such a block cipher around 2^{50} times and that can store and lookup around 2^{50} keys and messages. [6 marks]

(*c*) Your colleague has proposed the following digital signature algorithm. Let $(\mathbb{G}, q, g)$ be system-wide choices of a cyclic group $\mathbb{G}$ of prime order q with generator g such that the discrete logarithm problem in $\mathbb{G}$ is computationally infeasible. Further let $H : \{0,1\}^* \to \mathbb{Z}_q^*$ be a collision-resistant hash function. Pick a secret key $x \in \mathbb{Z}_q$ uniformly at random and let (y, r) with $y := g^x \in \mathbb{G}$ and $r := H(g^{H(x)})$ be the corresponding public key.

Then use as the signature of message $m \in \{0,1\}^*$ the value $s \in \mathbb{Z}_q^*$ found by solving

$$H(x) \cdot s \equiv x \cdot r + H(m) \pmod{q}$$

for $s = [H(x)]^{-1} \cdot [x \cdot r + H(m)]$. (Here a^{-1} denotes the multiplicative inverse of finite-field element $a \in \mathbb{Z}_q^*$. Your colleague considers $\mathbb{P}(s = 0)$ negligible.)

The recipient, given $(\mathbb{G}, q, g, H), (y, r), (m, s)$ verifies that signature by checking the equation

$$H\left(y^{r \cdot s^{-1}} g^{H(m) \cdot s^{-1}}\right) = r$$

Show that this signature scheme does not provide existential unforgability. [8 marks]

6 Denotational Semantics

(*a*) For a poset $(P, \sqsubseteq)$, the *join* of $x, y \in P$ is defined to be the element $x \sqcup y \in P$ such that

$$\forall\, p \in P.\; x \sqcup y \sqsubseteq p \iff (\, x \sqsubseteq p \wedge y \sqsubseteq p \,)$$

A poset is said to be *join complete* if every pair of elements in it has a join.

For a join-complete cpo D, show that the function $\sqcup : D \times D \to D$ mapping $(x, y) \in D \times D$ to $x \sqcup y \in D$ is continuous. [8 marks]

(*b*) Let $(D, \sqsubseteq)$ be a domain.

(*i*) For a continuous function $f : D \to D$, prove that the subset of D

$$\widehat{f} \;=\; \{\, x \in D \mid f(x) \sqsubseteq x \,\}$$

ordered by $\sqsubseteq$ is a domain. [6 marks]

(*ii*) For $d \in D$, let $\uparrow(d) = \{\, x \in D \mid d \sqsubseteq x \,\}$.

For a continuous function $g : D \to D$, prove that if $(D, \sqsubseteq)$ is join complete then, for all $d \in D$, the subset of D

$$\uparrow(d) \cap \widehat{g}$$

ordered by $\sqsubseteq$ is a domain. [6 marks]

7 E-Commerce

(*a*) Describe how a block chain distributed ledger operates and briefly describe two advantages and two disadvantages of such systems. [6 marks]

(*b*) Given that a number of systems are starting to use distributed ledgers to store and process transaction data, discuss the potential issues of using such ledgers given the rights data subjects have under GDPR. [8 marks]

(*c*) Discuss three network externalities or effects that could impact the adoption of a new distributed ledger technology for purchasing and tracking carbon emission offsets. [6 marks]

8 Hoare Logic and Model Checking

Consider commands C composed from assignments $X := E$ (where X is a program variable, and E is an arithmetic expression), heap allocation $X := \texttt{alloc}(E_1, ..., E_n)$, heap assignment $[E_1] := E_2$, heap dereference $X := [E]$, disposal of heap locations $\texttt{dispose}(E)$, the no-op `skip`, sequencing $C_1; C_2$, conditionals $\texttt{if } B \texttt{ then } C_1 \texttt{ else } C_2$ (where B is a boolean expression), and loops $\texttt{while } B \texttt{ do } C$. `null` is 0.
Recall the separation logic partial list representation predicates:

$$\begin{array}{lcl} \mathsf{plist}(t, [], u) & = & (t = u) \wedge \mathsf{emp} \\ \mathsf{plist}(t, h :: \alpha, u) & = & \exists y.\, ((t \mapsto h) * ((t+1) \mapsto y) * \mathsf{plist}(y, \alpha, u)) \end{array}$$

Circular lists can be represented by $\mathsf{clist}(t, \alpha) = \mathsf{plist}(t, \alpha, t) \wedge (\alpha = [] \Rightarrow t = \texttt{null})$.

(*a*) Assuming $\vdash \{P_1\}\ C_1\ \{Q_1\}$ and $\vdash \{P_2\}\ C_2\ \{Q_2\}$:

(*i*) explain precisely why $\vdash \{P_1 * P_2\}\ C_1; C_2\ \{Q_1 * Q_2\}$ [2 marks]

(*ii*) give a counterexample to $\vdash \{P_1 \wedge P_2\}\ C_1; C_2\ \{Q_1 \wedge Q_2\}$. [1 mark]

(*b*) Give a proof outline for the following circular list 'next' triple:

$$\{\mathsf{clist}(X, t :: \alpha)\}\ X := [X+1]\ \{\mathsf{clist}(X, \alpha +\!\!+ [t])\}$$ [3 marks]

(*c*) Give a loop invariant (no need for a proof outline) for the following circular list 'length' triple:]

$\{\mathsf{clist}(X, \alpha)\}$
$\texttt{if } X = \texttt{null } \texttt{then } Y := 0$
$\texttt{else } (Z := [X+1]; Y := 1; \texttt{while } Z \neq X \texttt{ do } (Z := [Z+1]; Y := Y+1))$
$\{\mathsf{clist}(X, \alpha) * Y = \mathit{length}(\alpha)\}$

[3 marks]

(*d*) Give a loop invariant (no need for a proof outline) for the following triple for a 'previous' operation on non-empty circular lists:

$\{\mathsf{clist}(X, \alpha +\!\!+ [t])\}$
$Z := X; Y := [X+1]; (\texttt{while } Y \neq X \texttt{ do } (Z := Y; Y := [Y+1])); X := Z$
$\{\mathsf{clist}(X, t :: \alpha)\}$

[4 marks]

(*e*) Give a loop invariant (no need for a proof outline) for the following triple for a 'dial to minimum' operation on non-empty circular lists:

$\{\mathsf{clist}(X, \alpha_1 +\!\!+ [t] +\!\!+ \alpha_2) \wedge \mathit{sorted}(t :: \mathit{merge}(\mathit{sort}(\alpha_1), \mathit{sort}(\alpha_2)))\}$
$Z := X; M := [X]; Y := [X+1];$
$(\texttt{while } Y \neq Z \texttt{ do}$
$\quad (N := [Y]; (\texttt{if } N < M \texttt{ then } X := Y \texttt{ else skip}); Y := [Y+1]));$
$\{\mathsf{clist}(X, [t] +\!\!+ \alpha_2 +\!\!+ \mathit{reverse}(\alpha_1))\}$

[5 marks]

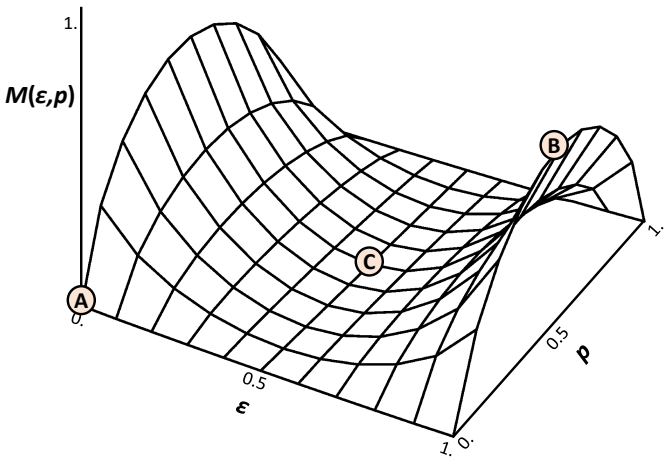
(*f*) Describe precisely *all* pairs of a stack and a heap that satisfy

$$\exists y, z.\, ((X \mapsto y * y \mapsto z * z \mapsto X) \wedge Y = 0)$$

[2 marks]

9 Information Theory

(*a*) A binary symmetric channel receives as input a bit whose values $\{0, 1\}$ have probabilities $\{p, 1-p\}$, but in either case, a transmission error can occur with probability ϵ which flips the bit. The surface plot below describes the mutual information of this channel as a function $M(\epsilon, p)$ of these probabilities:



(*i*) At the point marked A, the error probability is $\epsilon = 0$. Why then is the channel mutual information minimal in this case: $M(\epsilon, p) = 0$? [2 marks]

(*ii*) At the point marked B, an error always occurs ($\epsilon = 1$). Why then is the channel mutual information maximal in this case: $M(\epsilon, p) = 1$? [2 marks]

(*iii*) At the point marked C, the input bit values are equiprobable ($p = 0.5$), so the symbol source has maximal entropy. Why then is the channel mutual information in this case $M(\epsilon, p) = 0$? [2 marks]

(*iv*) Define mathematically the function $M(\epsilon, p)$ in terms of ϵ and p. [4 marks]

(*b*) An important operation in pattern recognition is convolution. If f and g are two functions $\mathbb{R} \to \mathbb{C}$ then their convolution is $(f * g)(x) = \displaystyle\int_{-\infty}^{\infty} f(x-y)g(y)dy$.

If their respective Fourier transforms are $\mathcal{F}_{[f]}(\omega)$ and $\mathcal{F}_{[g]}(\omega)$, prove that the convolution $(f*g)(x)$ has a Fourier transform $\mathcal{F}_{[f*g]}(\omega)$ that is the simple product

$$\mathcal{F}_{[f*g]}(\omega) = 2\pi\mathcal{F}_{[f]}(\omega) \cdot \mathcal{F}_{[g]}(\omega)\,.$$

[6 marks]

(*c*) Show how a generating (or "mother") wavelet $\Psi(x)$ can spawn a self-similar family of "daughter" wavelets $\Psi_{jk}(x)$ by simple scaling and shifting operations. Explain the advantages of analysing data in terms of such a self-similar family of dilates and translates of a mother wavelet. [4 marks]

10 Machine Learning and Bayesian Inference

(*a*) Give a detailed description of the general *Bayes decision rule* for classification. Include in your answer definitions of the *loss*, *conditional risk*, *decision rule* and *risk*. [7 marks]

(*b*) For a problem with C classes, we suffer a loss of 1 for an incorrect classification and 0 for a correct one. Show that the Bayes decision rule for inputs $\mathbf{x}$ is

$$h(\mathbf{x}) = \text{argmax}_c \Pr(c|\mathbf{x}).$$

[3 marks]

(*c*) For a problem with 2 classes, we now have three possibilities: classify $\mathbf{x}$ as being in class c_1, classify $\mathbf{x}$ as being in class c_2, or decline to classify $\mathbf{x}$. Classifying some $\mathbf{x}$ correctly results in a loss of 0 and classifying it incorrectly results in a loss of 1. Declining to classify $\mathbf{x}$ has a cost of θ_1 if $\mathbf{x}$ should be classified in class c_1 and θ_2 if it should be classified in class c_2. Both θ_1 and θ_2 can take values between 0 and 1/2.

Give a graphical representation of the conditional risks and use it to show that the Bayes decision rule for this problem is:

$$h(\mathbf{x}) = \begin{cases} c_1 & \text{if } p \leq q_1 \\ \text{Decline} & \text{if } q_1 < p \leq q_2 \\ c_2 & \text{if } q_2 < p \end{cases}$$

where $p = \Pr(c_1|\mathbf{x})$, $q_1 = \frac{\theta_2}{1-(\theta_1-\theta_2)}$ and $q_2 = \frac{1-\theta_2}{1+(\theta_1-\theta_2)}$. [10 marks]

11 Mobile and Sensor Systems

HearMe Ltd. has put a new wearable for the ear (earable) on the market that, in addition to other sensors, has an accelerometer. The company is hoping to use the accelerometer to analyze human activity recognition (HAR) of the users.

(*a*) What are the challenges the company will face in collecting the accelerometer data for human activity recognition (on their servers)? [5 marks]

(*b*) What are the steps of the HAR pipeline (you should also discuss person dependent/independent aspects)? Please use examples only related to accelerometer sensing.

[5 marks]

(*c*) What are the limitations of applying Deep Learning to HAR data for analysis and how can they be obviated?

[5 marks]

(*d*) The company would like to be able to perform HAR on-device on the earable. However, the current model running on the servers does not fit in its processor. What techniques can be applied to the model to make it run on the device efficiently and indicate in which way each helps? [5 marks]

12 Optimising Compilers

(*a*) Describe inference-based program analysis on expressions e. Explain how it can be used to judge effect systems. [3 marks]

(*b*) Give inference rules and a set of effects for an effect system for the following language:

$$e ::= x \mid \lambda x.e \mid e_1 e_2 \mid \texttt{if } e_1 \texttt{ then } e_2 \texttt{ else } e_3 \mid \mathcal{O}(\xi).e \mid \mathcal{W}(\xi e_1).e_2 \mid \mathcal{C}(\xi).e$$

where $\mathcal{O}(\xi).e$ opens file ξ and results in the value of e, $\mathcal{W}(\xi e_1).e_2$ evaluates e_1 and writes its (integer) value to file ξ before resulting in the value of e_2, and $\mathcal{C}(\xi).e$ closes file ξ before resulting in the value of e. The language types are integers and functions. [4 marks]

(*c*) Give and explain the safety condition for this system. [3 marks]

(*d*) Show how the rules from Part (*b*) assign a type and effect(s) to the following expression.

$$\mathcal{O}(\xi).\texttt{if } x \texttt{ then } \lambda x.\mathcal{W}(\xi x).\mathcal{C}(\xi).x \texttt{ else } \lambda x.\mathcal{C}(\xi).x$$

[5 marks]

(*e*) Justify that the effect system will allow us to identify all expressions with incorrect use of files or, using an example, describe why it won't and show how to alter the effect system so that it will allow us to. [5 marks]

13 Principles of Communications

(*a*) Briefly describe link-state routing, outlining the link-state advertisement and the shortest path computation that each node carries out. [5 marks]

(*b*) *Fibbing* is a technique for injecting link-state advertisements to the routing system that provides a way to add alternate paths into the results of the route computation. Describe what this is for, how this operates and what security implications it has.

[15 marks]

14 Quantum Computing

(a) In the superdense coding protocol, Alice and Bob each have one qubit from the entangled pair $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, which enables Alice to send two bits of classical information to Bob, using a single qubit. Explain the superdense coding protocol in detail, and show that it does indeed enable the transmission of two bits using a single qubit. [10 marks]

(b) This question concerns entangled states.

(i) Suppose Alice transmits the two-bit string '00' using the superdense coding protocol and an evesdropper, Eve, intercepts the qubit transmitted by Alice, measures it in the computational basis and then re-transmits to Bob. Find the probability that Bob correctly receives '00'. [5 marks]

(ii) Suppose further that prior to Eve's interception there is a 50% probability that the qubit experiences a bit-flip. What is the probability that Bob correctly receives 00? (Note that after Eve's retransmission no errors occur) [5 marks]

15 Types

(*a*) Recall that in constructive logic, logical negation is defined using implication and falsehood as $\neg A \triangleq A \supset \bot$.

(*i*) Does $A \supset \neg\neg A$? If so, give a simply-typed lambda-term corresponding to this implication.

(*ii*) Does $\neg\neg A \supset A$? If so, give a simply-typed lambda-term corresponding to this implication.

(*iii*) Does $\neg\neg\neg A \supset \neg A$? If so, give a simply-typed lambda-term corresponding to this implication.

[5 marks]

(*b*) (*i*) Give the typing rules for Peano natural numbers and their eliminator. [2 marks]

(*ii*) Using the rules given above, define the addition function. [3 marks]

(*iii*) Let a binary tree be either a leaf `Leaf` or a node `Node(l,x,r)` where `l` and `r` are subtrees, and `x` is a natural number. Give typing rules for trees corresponding to this prose description, including an eliminator. [3 marks]

(*iv*) Using the rules given above, define a function `size` which takes a binary tree and returns the total number of nodes in the tree. [5 marks]

(*c*) The `zip` function takes two lists, and returns a list of pairs of the elements as output. Suppose we see the following Agda type declaration for `zip`:

$$zip : \forall\{A\,B : Set\} \to \{n : Nat\} \to \mathsf{Vec}\,A\,n \to \mathsf{Vec}\,B\,n \to \mathsf{Vec}\,(A \times B)\,n$$

Explain what this means in terms of how to call the function, and what properties the result has. [2 marks]

END OF PAPER