COMPUTER SCIENCE TRIPOS Part II

Tuesday 4 June 2019 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*) For each of the following claims, state whether it is true or not and give a brief justification.

(*i*) For any linear program with n variables and m constraints, there are at most $\binom{n+m}{m}$ different basic solutions. [2 marks]

(*ii*) The Simplex Algorithm has a worst-case polynomial runtime. [2 marks]

(*iii*) In each iteration of the Simplex Algorithm, the value of the objective function changes. [2 marks]

(*iv*) The auxiliary linear program in Initialize-Simplex always has a feasible solution. [2 marks]

(*v*) The fundamental theorem of linear programming also holds if linear constraints are allowed to be strict. [2 marks]

(*vi*) The set of feasible solutions of any linear program forms a convex set. [2 marks]

(*b*) For the following linear program, write down the auxiliary linear program used by Initialize-Simplex in slack form: [3 marks]

$$\begin{aligned} \text{minimize} \quad & -4x_1 + x_2 \\ \text{subject to} \quad & -4x_1 + 2x_2 \geq -4 \\ & x_1 - 6x_2 \leq -3 \end{aligned}$$

(*c*) Recall the algorithm for the unweighted vertex cover problem that is based on rounding the solution of a linear program.

(*i*) What is the approximation ratio of this algorithm? [1 mark]

(*ii*) Give an example of a graph and the corresponding linear program for which the gap between the linear program solution and optimal solution is as large as possible. [4 marks]

2 Bioinformatics

(*a*) Describe Ruth Nussinov's algorithm on RNA folding and its complexity. Illustrate with one example. [4 marks]

(*b*) Describe the neighbour joining algorithm for phylogenetic analysis and its complexity. [5 marks]

(*c*) Hidden Markov models (HMM) are widely used in Bioinformatics.

(*i*) Describe how to build an HMM to identify exons and introns in genome sequences. [5 marks]

(*ii*) Discuss how to assess the performance of an HMM to identify exons and introns in genome sequences. [2 marks]

(*d*) Discuss the advantages and disadvantages of Leonard Adleman's approach to the travelling salesman problem with respect to the computational approach. [4 marks]

3 Comparative Architectures

(*a*) Describe the organisation of a two-level branch predictor that makes use of a global branch history. [3 marks]

(*b*) What are false (or name) dependencies and what hardware technique is often used to remove them within a processor? [3 marks]

(*c*) You are asked to design a new Instruction Set Architecture (ISA). You are told that this new ISA will be the basis for both simple low-power and high-performance processor implementations. Would you include the features or design choices listed below? In each case, carefully justify your answer.

(*i*) A branch delay slot.

(*ii*) The ability to predicate the execution of most instructions.

(*iii*) A conditional move instruction.

(*iv*) The use of condition codes (or flags) to specify the branch condition.

(*v*) A large number of general-purpose registers.

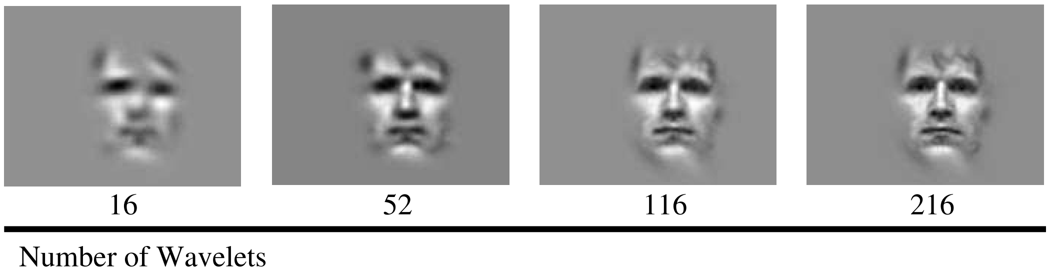
(*vi*) A TSO memory consistency model.

(*vii*) Support for custom ISA extensions, e.g. to allow the addition of special functional units.

[14 marks]

4 Computer Vision

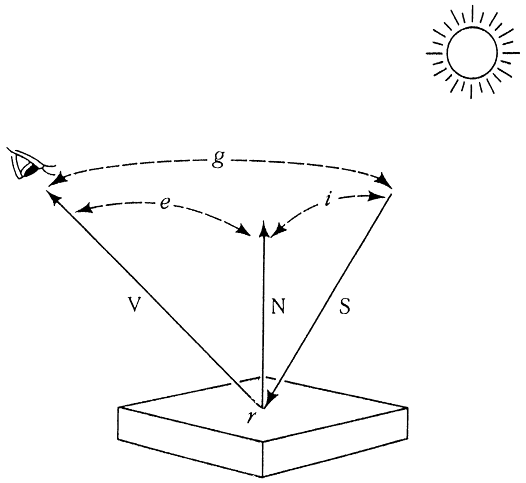
(*a*) Explain why such a tiny number of 2D Gabor wavelets as shown in this sequence are so efficient at representing faces, and why such wavelet-based encodings deliver good performance in “appearance-based” algorithms for face recognition. What role do such encodings have in representing different facial expressions? What sort of neural evidence is there for such encodings in human vision?



[6 marks]

(*b*) Explain the “receptive field” concept as used both in CNNs (convolutional neural networks) and in visual neuroscience, and explain the role of trainable connections. Why is the concept of convolution relevant? Roughly how many layers are in the ‘FaceNet’ CNN, and how many connection parameters must be trained? By comparison, in the visual cortex of the brain, typically how many synapses are there per neurone, and what is the total length of “wiring” (neurite connections) per cubic-millimeter? Can connection updates be a basis for computer vision? [6 marks]

(*c*) In relation to the image formation diagram shown below, 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*. 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. [8 marks]



5 Cryptography

(*a*) The *Tripos Encryption Standard (TES)* is a block cipher optimized for use on UGPs ("undergraduate processors"). It operates on 4-bit blocks, written as hexadecimal digits (e.g., a $\oplus$ 9 = 3). For one particular key K, it implements the following permutation:

m	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$E_K(m)$	1	b	5	c	7	e	2	a	4	9	f	d	0	3	6	8

Using this key K, decrypt the following three ciphertexts according to the indicated modes of operation. [*Note:* the XOR table at the bottom of this page may be of use.]

(*i*) ECB: 188b06 [2 marks]

(*ii*) CBC: 301b2 [3 marks]

(*iii*) CFB: 10f6d [3 marks]

(*b*) State four advantages that counter mode has over either CBC or CFB mode. [4 marks]

(*c*) Using the same K as in Part (*a*):

(*i*) Show that the CBC-MAC tag for message 1234 is d. [3 marks]

(*ii*) Demonstrate that CBC-MAC with a given K is not collision resistant, by showing how to find another message, of the form 1x04, that results in the same CBC-MAC message tag (without iterating over different candidates for 4-bit block x). [5 marks]

UGP XOR accelerator:

$\oplus$	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
1	1	0	3	2	5	4	7	6	9	8	b	a	d	c	f	e
2	2	3	0	1	6	7	4	5	a	b	8	9	e	f	c	d
3	3	2	1	0	7	6	5	4	b	a	9	8	f	e	d	c
4	4	5	6	7	0	1	2	3	c	d	e	f	8	9	a	b
5	5	4	7	6	1	0	3	2	d	c	f	e	9	8	b	a
6	6	7	4	5	2	3	0	1	e	f	c	d	a	b	8	9
7	7	6	5	4	3	2	1	0	f	e	d	c	b	a	9	8
8	8	9	a	b	c	d	e	f	0	1	2	3	4	5	6	7
9	9	8	b	a	d	c	f	e	1	0	3	2	5	4	7	6
a	a	b	8	9	e	f	c	d	2	3	0	1	6	7	4	5
b	b	a	9	8	f	e	d	c	3	2	1	0	7	6	5	4
c	c	d	e	f	8	9	a	b	4	5	6	7	0	1	2	3
d	d	c	f	e	9	8	b	a	5	4	7	6	1	0	3	2
e	e	f	c	d	a	b	8	9	6	7	4	5	2	3	0	1
f	f	e	d	c	b	a	9	8	7	6	5	4	3	2	1	0

6 E-Commerce

(*a*) Describe five ecommerce business models. [5 marks]

(*b*) Describe five things to consider when internationalising an ecommerce business. [5 marks]

(*c*) Does the nature of the Internet naturally lead to dominant firms in ecommerce markets? [10 marks]

7 Hoare Logic and Model Checking

Consider a programming language that consists of commands C composed from assignments $X := E$ (where X is a program variable, and E is an arithmetic expression), heap allocation $X := \texttt{alloc}(E_1, \ldots, E_n)$, heap assignment $[E_1] := E_2$, heap dereference $X := [E]$, disposal of heap locations $\texttt{dispose}(E)$, the no-op $\texttt{skip}$, 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 $\texttt{while}\ B\ \texttt{do}\ C$. $\texttt{null}$ is 0

(*a*) Explain informally what it means for a separation logic partial correctness triple $\{P\}\ C\ \{Q\}$ to be valid. [3 marks]

(*b*) Explain informally what it means in terms of the executions of C for the separation logic partial correctness triple $\{\top\}\ C\ \{\bot\}$ to be valid. [2 marks]

(*c*) Recall the list representation predicate *list*:

$$\mathit{list}(t, []) = (t = \texttt{null}) \quad \mathit{list}(t, h :: \alpha) = \exists y.\, ((t \mapsto h) * ((t+1) \mapsto y) * \mathit{list}(y, \alpha))$$

We write $[]$ for the empty mathematical list; $h :: \alpha$ for the mathematical list the head of which is h, and the tail of which is α; $\alpha +\!\!+\ \beta$ for the concatenation of mathematical lists α and β; $\alpha[i]$ for the i-th element of the list α, starting at 0; and $[k, \ldots, n]$ for the ascending list of integers from k to n, including k and n. Give a proof outline, including a loop invariant, for the following triple:

$\{N = n \wedge N \geq 0\}$
$X := \texttt{null}; \texttt{while}\ N > 0\ \texttt{do}\ (X := \texttt{alloc}(N, X); N := N - 1)$
$\{\mathit{list}(X, [1, \ldots, n])\}$ [4 marks]

(*d*) Also recall the partial list representation predicate *plist*:

$$\mathit{plist}(t, [], u) = (t = u)$$
$$\mathit{plist}(t, h :: \alpha, u) = \exists y.\, ((t \mapsto h) * ((t+1) \mapsto y) * \mathit{plist}(y, \alpha, u))$$

Give a loop invariant for the following list sum triple:

$\{\mathit{list}(X, \alpha)\}$
$Y := X; N := 0; \texttt{while}\ Y \neq \texttt{null}\ \texttt{do}\ (M := [Y]; N := N + M; Y := [Y + 1])$
$\left\{\mathit{list}(X, \alpha) \wedge N = \sum_{i=0}^{\mathit{length}(\alpha)-1} \alpha[i]\right\}$ [4 marks]

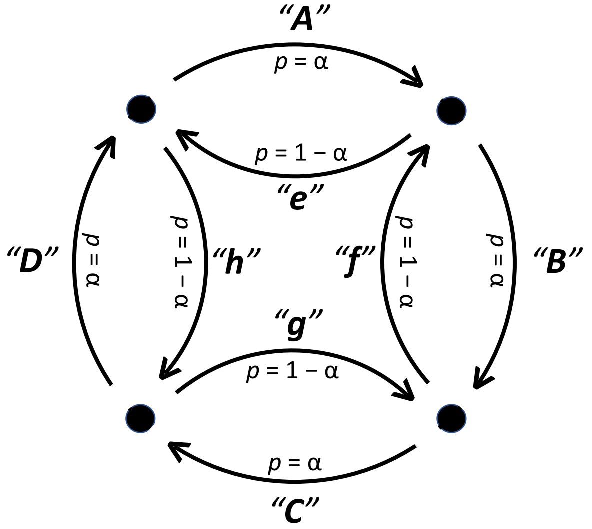
(*e*) Give a loop invariant for the following list concatenation triple:

$\{\mathit{list}(X, \alpha) * \mathit{list}(Y, \beta)\}$
$\texttt{if}\ X = \texttt{null}\ \texttt{then}\ Z := Y\ \texttt{else}$
$$\left(\begin{array}{l} Z := X; U := Z; V := [Z+1]; \\ \texttt{while}\ V \neq \texttt{null}\ \texttt{do}\ (U := V; V := [V+1]); \\ [U+1] := Y \end{array}\right)$$
$\{\mathit{list}(Z, \alpha +\!\!+\ \beta)\}$ [5 marks]

(*f*) Describe precisely a stack and a heap that satisfy $\mathit{list}(X, [1, \ldots, 3])$. [2 marks]

8 Information Theory

(*a*) Consider the four-state Markov process in the graph below. It emits the eight letters $\{A, B, C, D, e, f, g, h\}$ with probabilities and changes of state as shown, but note the sequence constraints. (For example, an A can only be followed by a B or an e.) Letter emissions with clockwise state transitions occur with probability α, and the others with probability $1 - \alpha$, where $0 < \alpha < 1$.



(*i*) First imagine a one-state Markov process that emits any of eight letters with equal probabilities. What is its entropy? [2 marks]

(*ii*) For the four-state Markov process shown with parameter α, what is the long-term probability distribution across the eight letters? [4 marks]

(*iii*) In terms of parameter α, what is the overall entropy $H(\alpha)$ of this four-state Markov process? [2 marks]

(*iv*) Sketch a plot of $H(\alpha)$ as a function of α. Compare its maximum value with your earlier answer in Part (*a*)(*i*) for a one-state Markov process that also emits eight letters, and explain the difference, if any. [4 marks]

(*b*) Is it possible to construct an instantaneous code (a code possessing the prefix property) for a five-letter symbol set using codewords whose lengths in bits are: 1, 2, 3, 3, and 4 bits? Justify your answer by stating the relevant condition. [4 marks]

(*c*) Provide an operation in linear algebra that involves simply the multiplication of a matrix by a vector, which describes the Discrete Fourier Transform of a discrete sequence of data $f[n] = (f[1], ..., f[N])$ to obtain Fourier coefficients $F[k] = (F[1], ..., F[N])$. Define the elements of the $(N \times N)$ matrix and give the computational cost of the operation in this vector-matrix form. [4 marks]

9 Machine Learning and Bayesian Inference

In designing a system to perform linear regression with noisy data, you feel that the noise is not modelled well by the usual normal density, and wish instead to use the *Cauchy density*

$$p(x) = \frac{1}{\beta\pi}\left(\frac{\beta^2}{(x-\alpha)^2+\beta^2}\right)$$

having parameters α and β.

(*a*) Denote the weights of your model by the vector $\mathbf{w}$. Given a set $\mathbf{s}$ of m examples, each consisting of a d-dimensional vector $\mathbf{x}$ and corresponding label y, find an expression for the *likelihood* $p(\mathbf{y}|\mathbf{w})$ where $\mathbf{y}^T = (y_1, \dots, y_m)$. State any assumptions you make. [6 marks]

(*b*) In addition to the unusual noise density, you have some knowledge of the problem at hand suggesting that some of the parameters in $\mathbf{w}$ are likely to be close to known values, whereas the others have no such constraint. Suggest a suitable *prior density* $p(\mathbf{w})$ that could be used to model this. You should assume that the weights are independent for the purposes of designing a prior, and you may use the fact that the normal density is

$$p(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

[5 marks]

(*c*) Using your answers to parts (*a*) and (*b*) derive a *maximum a posteriori (MAP)* learning algorithm for the problem. Should your algorithm require derivatives you may state them without working them out in full. [6 marks]

(*d*) Suggest a way in which any parameters other than $\mathbf{w}$ might be either given an effective value or removed from consideration. Give a single advantage and a single disadvantage of the method you suggest. [3 marks]

10 Mobile and Sensor Systems

A team of geologists want to set up a system to monitor an area close to a dormant volcano for seismic movements. They have a single basestation which will connect to cellular infrastructure and transmit the data back to their server. They plan to scatter sensor nodes in the area to be monitored.

(*a*) The team has decided to use SMAC for their Medium Access Control (MAC) layer routing protocol. Explain the advantages and limitations of using SMAC in such a system. [5 marks]

(*b*) Illustrate a multihop routing solution. Describe in detail the protocol you would use for this system indicating advantages and limitations. [4 marks]

(*c*) Describe a single hop solution using an IoT protocol and explain its strengths and weaknesses (include details of the protocol). Discuss whether this solution is better or worse than the solution in Part (*b*). [5 marks]

(*d*) Now assume that the sensor nodes are mobile (e.g., deployed on autonomous mobile ground platforms).

(*i*) Explain the benefits of using autonomous mobile sensor nodes in this application. [2 marks]

(*ii*) Describe a mechanism by which the mobile sensor nodes can collaborate to improve the monitoring information obtained. Include details of your estimation framework, and explain how it would be used to inform the motion planning of your mobile sensor nodes. [4 marks]

11 Optimising Compilers

The following C-style code from an untyped language is analysed by a compiler, where the `work()` function is assumed to have no side effects.

```
1   c = &b;
2   *c = &c;
3   a = c;
4   c = &d;
5   if (v == 0)
6     *c = **a;
7   else
8     *c = *b;
9   *a = &a;
10  work(a);
11  work(c);
```

(*a*) Describe alias analysis and the transformations it enables. [4 marks]

(*b*) Summarise Andersen's analysis and calculate the points-to set, $pt(x)$, for each pointer, x, within the C-style code above. [9 marks]

(*c*) Describe the reason that the analysis overestimates some of the sets in the answer to Part (*b*). [2 marks]

(*d*) Now assume that the `work()` function may alter memory locations reachable through its argument. Explain why the two calls to `work()` in lines 10 and 11 cannot be executed concurrently using the analysis from Part (*b*), but can be based on the answer to Part (*c*). [5 marks]

12 Principles of Communications

(*a*) The Border Gateway Protocol (BGP) uses attributes to enforce transit relationships for Outbound route filtering, and to enforce the order of route preference between customers, peers and providers.

Key attributes, in precedence order, are as follows (from highest to lowest):

- Highest Local Preference
- Shortest ASPATH
- Lowest MED
- i-BGP < e-BGP
- Lowest IGP cost to BGP egress
- Lowest router ID

Discuss the basic use of the relevant BGP mechanisms and their use of attributes for backup and for traffic engineering.

(*i*) Explain the different attributes.

(*ii*) Explain the mechanisms that are used to determine paths within an AS.

(*iii*) Explain the mechanisms that are used to determine paths between ASs.

[5 marks each]

(*b*) BGP announces and withdraws prefixes between Autonomous Systems so that different domains can route traffic amongst themselves. The dynamics of BGP advertisements can be impacted by intra-domain routing, so that if a route flaps, for example because of an intermittent fault on a link or router, this can be exported to the whole Internet. Simple techniques of fixing an interval for advertisements, and *punishing* routers that exceed that rate have been proposed.

Describe very briefly how one might use a control theoretic approach to provide stable damping of this effect dynamically, and possibly, more efficiently.

[5 marks]

13 Types

Recall the three judgements for classical propositional logic:

(*a*) $\Gamma; \Delta \vdash e : A\ \mathsf{true}$ – e is a proof of type A

(*b*) $\Gamma; \Delta \vdash k : A\ \mathsf{false}$ – k is a refutation of type A

(*c*) $\Gamma; \Delta \vdash \langle e \mid_A k\rangle\ \mathsf{contr}$ – $\langle e \mid_A k\rangle$ is a contradiction at type A

Here, Γ contains the true assumptions, and Δ are the false assumptions. In this question, we will extend classical propositional logic with support for the implication or function type operator $A \to B$.

(*a*) Give a proof term and inference rule for a proof of type $A \to B$. [4 marks]

(*b*) Give a proof term and inference rule for a refutation of type $A \to B$.

[*Hint*: how is implication encoded in classical logic?] [4 marks]

(*c*) Give a reduction rule for contradiction configurations of the form $\langle e \mid_{A\to B} k\rangle$. [4 marks]

(*d*) (*i*) State the preservation theorem for classical logic. [2 marks]

(*ii*) Give the proof of preservation for the case of the new rule defined above. You may assume that weakening, exchange and substitution all hold. [6 marks]

END OF PAPER