COMPUTER SCIENCE TRIPOS Part II

Wednesday 6 June 2018 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*) The rendering equation is often formulated as:

$$L_r(\boldsymbol{\omega_r}) = \int_{\Omega} \rho(\boldsymbol{\omega_i}, \boldsymbol{\omega_r})\, L_i(\boldsymbol{\omega_i}) \cos\theta_i \, d\boldsymbol{\omega_i} \tag{1}$$

Briefly explain each term in this equation, including integration domain. Support your answer with a drawing. [7 marks]

(*b*) Explain why the rendering equation is computationally expensive to solve for complex scenes. [7 marks]

(*c*) Sometimes in mornings and evenings when the sun is low above the horizon and there is a thin layer of water on the road, the road surface reflects large quantities of light, introducing a strong source of glare that makes the road difficult to see. Explain what reflection properties of the road surface are responsible for this effect. Why can we not observe such reflection when the sun is high above the horizon? [6 marks]

2 Comparative Architectures

(*a*) Total Store Order (TSO) is a widely implemented memory consistency model. How does TSO differ from Sequential Consistency (SC)? [4 marks]

(*b*) Describe an execution of a simple program that is legal under TSO but not SC. The simple program should consist of two threads, each thread consisting of a small number of load and store instructions accessing memory locations X and Y. [3 marks]

(*c*) Would the use of a coalescing write buffer violate TSO? Justify your answer. [4 marks]

(*d*) Consider a chip-multiprocessor where each core supports simultaneous multi-threading (SMT). Furthermore, each processor core has a write-through L1 data cache. Could multicopy atomicity be supported in such a design? [6 marks]

(*e*) Consider a chip-multiprocessor with two cores P1 and P2. Coherence is maintained using a MESI protocol with support for Bus Upgrade (`BusUpgr`) transactions to avoid unnecessary data transfers (i.e. when moving from state S to M). Why might a processor that initially makes a `BusUpgr` request have to change the request to a Read-Exclusive (`BusRdX`) before it even wins access to the shared bus? [3 marks]

3 Computer Systems Modelling

This question deals with stochastic processes $\{N(t), t \geq 0\}$ where $N(t)$ represents the number of events in the time interval $[0, t]$.

(a) (i) Define a Poisson process $\{N(t), t \geq 0\}$ of rate $\lambda > 0$. [2 marks]

(ii) Show that $N(t) \sim \text{Pois}(\lambda t)$ for each fixed $t > 0$. You may use the result that $\lim_{n\to\infty}(1 - x/n)^n = e^{-x}$ without proof. [4 marks]

(iii) Let X_1 be the time of the first event of the Poisson process $N(t)$. Show that $X_1 \sim \text{Exp}(\lambda)$. [2 marks]

(iv) Now given that $N(t) = 1$ derive the distribution of the time of the single event in $[0, t]$. [4 marks]

(b) Suppose that events of a Poisson process of rate λ are independently selected at random with probability $p > 0$. Show that the process of selected events is also a Poisson process and establish its rate. [2 marks]

(c) Describe how your result from part (b) can be used to simulate a non-homogeneous Poisson process whose rate function $\lambda(t)$ is such that $\lambda(t) \leq \lambda^*$ for all $t \geq 0$. [6 marks]

4 Computer Vision

(a) Explain how each of the following equations or expressions can be used for detecting and estimating visual motion in a spatio-temporal image sequence $I(x, y, t)$. Include in your answer the name used to describe each of these general classes of motion extraction models:

(i) $$-\frac{\partial I(x,y,t)}{\partial t} = \vec{v} \cdot \vec{\nabla} I(x,y,t)$$ [2 marks]

(ii) $$-\frac{\partial}{\partial t}\left[\nabla^2 G_\sigma(x,y) * I(x,y,t)\right]$$ [2 marks]

(iii) $$\text{argmax}_{(v_x,v_y)} \int_x \int_y \int_t I(x,y,t) \cdot I(x - v_x\tau,\ y - v_y\tau,\ t - \tau)\ dx\, dy\, dt$$ [2 marks]

(iv) $$F(\omega_x, \omega_y, \omega_t) = e^{-i(\omega_x v_x \tau\ +\ \omega_y v_y \tau\ +\ \omega_t \tau)} F(\omega_x, \omega_y, \omega_t)$$
where $$F(\omega_x, \omega_y, \omega_t) = \int_x \int_y \int_t I(x,y,t)\, e^{-i(\omega_x x + \omega_y y + \omega_t t)}\ dx\, dy\, dt$$ [2 marks]

(b) Colour perception is not about measuring wavelengths, because they vary with illumination. Explain why it is difficult to assign intrinsic spectral reflectance properties of surfaces. Explain all steps in the Retinex Algorithm intended to solve this, relating these steps where possible to neurobiology. [7 marks]

(c) Sketch out an algorithm for shape classification and the construction of shape grammars, involving active contours, codon strings, and indexing. Explain how codon constraints enable a shape grammar to define broad equivalence classes such as "cashew shaped" objects, with invariance to irrelevant transformations such as planar rotations or dilations. [5 marks]

5 Digital Signal Processing

Your friend Sam works on a physics experiment. This generates a voltage waveform $v(t)$ that is the sum of several signals:

- a sine wave $s(t) = A \cdot \sin(2\pi t f + \phi)$, the frequency f and phase ϕ of which are not known in advance, but f will be within 9.6 kHz $< f <$ 12.0 kHz;
- several other sine waves with frequencies below 8 kHz that Sam needs to ignore in her measurements;
- low levels of noise at all frequencies.

Sam needs to estimate the amplitude A of $s(t)$. She uses a USB audio recorder with a built-in 16 kHz anti-aliasing low-pass filter to digitize $v(t)$ at sampling frequency $f_\mathrm{s} = 48$ kHz, recording $s = 100\,000$ consecutive samples, resulting in real-valued samples $v_0, \ldots, v_{s-1}$. She implemented this algorithm to estimate A:

1: **input** $v_0, \ldots, v_{s-1}$

2: $b := 1000$; $\quad c := \left\lfloor \frac{s}{b} \right\rfloor$

3: $w_{k,l} := v_{kb+l}$ for all $0 \le k < c$, $0 \le l < b$

4: $x_{k,n} := \sum_{m=0}^{b-1} w_{k,m} \cdot \mathrm{e}^{-2\pi \mathrm{j} \frac{nm}{b}}$ for all $0 \le k < c$, $0 \le n < b$

5: $y_n := \left| \frac{1}{c} \cdot \sum_{k=0}^{c-1} x_{k,n} \right|$ for all $0 \le n < b$

6: $z := \max\{y_{n_1}, \ldots, y_{n_2}\}$ with $n_1 = 200$, $n_2 = 220$

7: **output** z

(*a*) Sam hopes that $A \approx z \cdot \alpha$ for some calibration constant α. She tries to determine α by connecting the USB audio recorder's input to a calibrated laboratory sine-wave generator set to output an amplitude of "60.0 dBμV". What amplitude A in volts will this test signal $A \cdot \sin(\ldots)$ have? [3 marks]

(*b*) When Sam varies the test-signal frequency f in the range 9.6–12.0 kHz, she is disappointed that the output z varies greatly: for some f it even drops to zero!

Describe what Sam's algorithm tries to do, identify and explain *three* problems in it, and change *three* lines to make z more proportional to A across the expected range of f, and close to zero outside that range. [15 marks]

(*c*) Suggest a small adjustment to b to accommodate a faster algorithm for one of the above steps. [2 marks]

6 E-Commerce

(*a*) Describe the aspects of a fair market. [4 marks]

(*b*) Describe the settlement mechanism in a credit card transaction. [4 marks]

(*c*) Describe the functions of a currency and the challenges of implementing electronic currencies. [4 marks]

(*d*) Discuss whether Bitcoin is better described as a currency, a bearer certificate or something else and what tests you would apply in making this determination. [8 marks]

7 Information Retrieval

(*a*) (*i*) Describe the difference between relevance feedback and query expansion in terms of user interaction. [2 marks]

(*ii*) Explain what we mean by equivalence classing of terms and why it is useful. Give one example of equivalence classing that can fail to retrieve the right documents. [2 marks]

(*iii*) Give an example of how asymmetric expansion of query terms can usefully model users' expectations. Is asymmetric expansion of query terms more or less efficient than equivalence classing? Justify your answer. [2 marks]

(*iv*) How would you evaluate an Information Retrieval task for which there is high tolerance for overlooked relevant information items? How might you modify the *F measure* for such a task? Justify your answer. [3 marks]

(*b*) (*i*) Given the query "**elvis music**" and the following term frequencies for the three documents *doc1*, *doc2* and *doc3*:

	elvis	**presley**	**mississippi**	**pop**	**music**	**life**
doc1	3	4	0	6	0	0
doc2	4	0	4	0	0	3
doc3	5	3	0	4	4	0

calculate the cosine similarity between the query and each document (you can ignore the *idf* term) in order to rank these documents in order of relevance. Show your workings. [2 marks]

(*ii*) The Rocchio algorithm is a classic algorithm for implementing relevance feedback. Use Rocchio to compute the new query vector for "**elvis music**" using *doc3* for relevance feedback (i.e., *doc3* has been marked as relevant). Give suitable values for Rocchio's weight parameters. As above, calculate cosine similarity (you can ignore the *idf* term) in order to rank the documents in order of relevance. Show your workings. [4 marks]

(*iii*) In Rocchio's algorithm, positive feedback is typically more useful than negative feedback. Give two example cases of negative documents being fed back that can decrease the retrieval effectiveness of the Rocchio re-formulated query. Propose one way in which you can incorporate negative feedback more effectively, and explain why this helps with the two examples you provided. Motivate your answer. [5 marks]

8 Machine Learning and Bayesian Inference

Evil Robot has decided to become a gambling cheat. He has a biased coin with $\Pr(\text{head}) = p$ and two dice. The first die is biased with $\Pr(n) = p_n$ for the nth outcome with $n \in \{1, 2, 3, 4, 5, 6\}$. The second die is also biased, and has different numbers: its distribution is $\Pr(n) = q_n$ with $n \in \{4, 5, 6, 7, 8, 9\}$.

Evil Robot flips the coin. If he gets a head then he rolls the first die, otherwise he rolls the second. He then tells you the outcome. You see only the number obtained and nothing else. He does this m times, so you observe a sequence of m numbers in the range 1 to 9. Your aim is to estimate p and the distributions of each die, given the m numbers. In the following, $\mathbf{n}$ is the vector of m observed numbers $(n_1 \quad \cdots \quad n_m)^T$, θ is the set of parameters $\{p, p_1, \ldots, p_6, q_4, \ldots, q_9\}$ and we define $q = 1 - p$.

(*a*) Write down an expression for the distribution $\Pr(n|\theta)$ where $n \in \{1, \ldots, 9\}$. [2 marks]

(*b*) Define the variable

$$z_i = \begin{cases} 1 & \text{if } n_i \text{ was obtained by rolling die 1} \\ 0 & \text{otherwise.} \end{cases}$$

and let $\mathbf{z}$ denote the corresponding vector with m values. Write down an expression for $\log \Pr(\mathbf{n}, \mathbf{z}|\theta)$. [3 marks]

(*c*) Describe the EM algorithm for maximizing likelihood in a problem involving *latent variables*. [3 marks]

(*d*) Show that, with the distribution $\Pr(\mathbf{z}|\mathbf{n}, \theta)$,

$$E(z_i) = \begin{cases} 1 & \text{if } n_i \in \{1, 2, 3\} \\ 0 & \text{if } n_i \in \{7, 8, 9\} \\ \frac{pp_{n_i}}{pp_{n_i} + qq_{n_i}} & \text{otherwise.} \end{cases}$$

[4 marks]

(*e*) Define $\gamma_i = E(z_i)$ as in Part (*d*). By applying the EM algorithm to this problem, show that you can estimate the parameters in θ using the following updates

$$p = \frac{\gamma}{m}$$
$$p_n = \frac{1}{\gamma} \sum_{\{i|n=n_i\}} \gamma_i$$
$$q_n = \frac{1}{m - \gamma} \sum_{\{i|n=n_i\}} (1 - \gamma_i)$$

where $\gamma = \sum_{i=1}^{m} \gamma_i$. [8 marks]

9 Mobile and Sensor Systems

A leisure park has decided to adopt a variety of mobile and sensing technologies to monitor the usage of its attractions. An app is offered to customers willing to install it on their smartphone, which tracks customers' location throughout the park and uses the phone accelerometer to monitor activity.

(*a*) Describe how app developers could make sure the app delivers the best accuracy on location and activity tracking while preserving the phone battery as much as possible through system and sensor sampling optimizations. [6 marks]

(*b*) A customer installs the app on their device. They have two other applications on the phone which monitor their physical activity as well as their location throughout the day for clinical reasons. Describe how the phone operating system could optimize the battery efficiency of the sensing across applications. [5 marks]

(*c*) A variety of wireless sensors are scattered throughout the park to monitor the operation of the attractions by continuously gathering temporal data (e.g., mechanical vibration, load, temperature, humidity).

(*i*) Describe a combination of medium access and network layer protocols for infrastructure-less multi-hop networks to aid the sensor data delivery to the park management servers. Discuss advantages and disadvantages of the solution devised. [5 marks]

(*ii*) Describe a solution which uses protocols from the Internet of Things domain to offer a non multi-hop solution. Discuss the advantages and disadvantages of the solution devised. [4 marks]

10 Principles of Communications

(*a*) There are two main ways in which traffic sources can be controlled: open-loop and closed-loop. Describe the overall operation and use of these two kinds of network system: What are the key components of an open-loop controlled protocol? What are the key components of a closed-loop, or feedback controlled system? What kinds of applications are suited to each class of control loop?
[10 marks]

(*b*) Two important approaches to fairness are max-min, and proportional. Describe what properties each of these offer, how they differ in how they can be used, and what happens in each case when demand exceeds supply of network capacity.
[10 marks]

11 Quantum Computing

(*a*) In the quantum teleportation protocol, Alice and Bob are each in possession of one qubit of a pair in the joint state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. In addition, Alice has a qubit in an arbitrary state $|\phi\rangle$. Explain how the protocol works. In particular, show that it involves the transmission of exactly two classical bits of information from Alice to Bob and demonstrate how, at the end of the protocol, Bob is in possession of a qubit in state $|\phi\rangle$. [10 marks]

(*b*) Suppose now that Alice has two qubits in a state

$$|\theta\rangle = \alpha_0|00\rangle + \alpha_1|01\rangle + \alpha_2|10\rangle + \alpha_3|11\rangle.$$

In addition, there is a pair of qubits in the state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ such that Alice is in possession of one qubit of the pair while Bob is in possession of the other.

Alice now uses the quantum teleportation protocol to transmit to Bob the first qubit of $|\theta\rangle$. What is the resulting joint state of the two-qubit system composed of the second of Alice's qubits and the qubit in Bob's possession? [10 marks]

12 Security II

(*a*) Name *three* different families of algebraic groups that are commonly used in cryptographic applications of the Diffie–Hellman problem, where *any* group element (other than the neutral element) can be used as a generator. Briefly outline some of their main attributes, such as the set of elements and the group operator. [9 marks]

(*b*) You are preparing to participate in a password-cracking competition. During the competition, you will be given the 128-bit hash-function output MD5(p). You have to find p, an 8-character password, each character having been chosen uniformly at random from a known alphabet of 64 ASCII characters.

In the weeks preparing for the competition, you have access to a small cluster of GPU graphics cards that can evaluate MD5 10^9 times per second.

During the competition, you have only access to a laptop computer that can evaluate MD5 10^6 times per second.

Without any pre-computation, how long would it take to evaluate MD5 for all possible passwords p in a brute-force attack

(*i*) on the laptop? [2 marks]

(*ii*) on the GPU cluster? [2 marks]

You decide to use the GPU cluster to pre-compute a *rainbow table* for this challenge.

(*iii*) What functions other than MD5 will the GPU cluster have to evaluate as often as MD5 when building the rainbow table? [3 marks]

(*iv*) Your laptop has enough RAM for storing the rainbow table as a hash table of 2^{32} key-value pairs (x, y) with $x, y \in \{0,1\}^{128}$. If you execute MD5 2^{50} times while generating your rainbow table, how long will your laptop need (worst case) to find a password p stored in it, given its MD5 hash value MD5(p)? Assume that the runtime is entirely dominated by the MD5 evaluations. [4 marks]

13 System-on-Chip Design

(*a*) At the lowest level, what is the primary consumer of electrical power in digital logic today? Give a formula for the expected energy or power use for a CMOS gate. [2 marks]

(*b*) A matrix (a 2-dimensional array) is stored on-chip in static RAM. What main factors contribute to the time and energy needed to transpose it? [4 marks]

(*c*) Assume now a square matrix is to be held in DRAM.

(*i*) When might it be helpful to store multiple-copies of a given matrix in different DRAM banks? [1 mark]

(*ii*) When might it be helpful to store multiple-copies of the matrix (or another example data structure) in one DRAM bank? [2 marks]

(*iii*) One way to avoid transposing a matrix is simply to hold an annotation that it has been transposed and to then swap over the row and column arguments for each operation. Why might physically performing the transpose ultimately benefit performance? Where would the annotation be held? [2 marks]

(*d*) A computation operates on square matrices of size $10^5 \times 10^5$. The inner loop, to be accelerated in hardware, has the following basic structure:

```
for (int i= ...) for (int j= ...)
{
   DD[i, j] = ff(SS[i-1, j], SS[i, j-1])
}
```

(*i*) Are there any loop-carried dependencies? What does this mean for performance optimisation? [1 mark]

(*ii*) If the DRAM timings are 11-11-11, meaning row activation, column activation and writeback each take 11 clock cycles, estimate roughly the minimum time for a naive implementation of the computation. Assume a simple linear data layout. State all further assumptions. [6 marks]

(*iii*) What determines whether it is possible or a good idea to perform the operation 'in place' (ie. using the same memory for DD and SS)? [2 marks]

14 Types

(*a*) Give typing rules for the introduction form $\mathsf{pack}(\tau, M)$ and elimination form $\mathsf{unpack}\ M\ \mathsf{as}\ (x, \alpha)\ \mathsf{in}\ N$ of the existential type $\exists\alpha(\tau)$.

[4 marks]

(*b*) An infinite stream of booleans can be represented in the polymorphic lambda calculus using the existential type

$$\mathsf{stream} \triangleq \exists\alpha(\alpha \times (\alpha \to (\mathsf{bool} \times \alpha)))$$

(*i*) Using the encoding above, define a function $\mathsf{head} : \mathsf{stream} \to \mathsf{bool}$.
[3 marks]

(*ii*) Using the encoding above, define a function $\mathsf{tail} : \mathsf{stream} \to \mathsf{stream}$.
[3 marks]

(*iii*) Using the encoding above, define a function

$$\mathsf{unfold} : \forall\alpha(\alpha \to (\alpha \to (\mathsf{bool} \times \alpha)) \to \mathsf{stream})$$

[4 marks]

(*iv*) Using unfold and the other functions above, define a function $\mathsf{notstream} : \mathsf{stream} \to \mathsf{stream}$, which returns a stream containing the boolean negation of the elements of the input stream. This answer should not use explicit pack or unpack expressions. [6 marks]

[*Note:* You may use extensions to the pure polymorphic lambda calculus such as let-bindings, natural numbers, products, and sum types, but carefully note their use and their typing in your answers.]

END OF PAPER