CST2
COMPUTER SCIENCE TRIPOS Part II

Tuesday 9 June 2026 14:00 to 17:00

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

SPECIAL REQUIREMENTS
Approved calculator permitted

1 Advanced Computer Architecture

(*a*) Imagine a multicore processor where no core has a private cache. Consider the following list of configurations which describes if each core has a write buffer and if so, how it operates. In each case, what is the strongest memory consistency model (SC, TSO or relaxed) that can be supported in a straightforward manner? Briefly justify each answer.

(*i*) No write buffer. [2 marks]

(*ii*) A private non-coalescing write buffer. If a read refers to an address that has a pending write in the buffer, the buffer is flushed. [2 marks]

(*iii*) A private non-coalescing write buffer where store values from the buffer can be forwarded to loads if required. [2 marks]

(*iv*) A private coalescing write buffer where store values from the buffer can be forwarded to loads if required. [2 marks]

(*b*) Consider a simple SoC built from a general-purpose CPU, caches, and off-chip I/O. List three ways in which this design could be extended or specialised to improve performance when running a particular program. In each case, discuss both the performance benefits for the target program and the potential disadvantages for other workloads. [9 marks]

(*c*) An optimising compiler or performance-conscious programmer may apply cache blocking (also known as tiling) to improve performance. How does the presence of a multi-level cache hierarchy (e.g. L1, L2 and L3 caches) affect the application of cache blocking? [3 marks]

2 Algebraic Techniques for Programming

(*a*) Recall that Datafun is the typed lambda calculus introduced to formulate fixed point computations, in which every type denotes a partially ordered set.

Suppose we want to extend Datafun with a new type constructor $X_\perp$, which adjoins a new element to X, with the property that the new element is below every element of X.

(*i*) Give an interpretation of the $X_\perp$ type constructor as a functor. [2 marks]

(*ii*) Give typing rules to introduce and eliminate this type. [4 marks]

(*iii*) Give a denotational interpretation of these rules, and explain why they are sound. [4 marks]

(*b*) A graph G consists of a finite set of nodes N, a finite set of edges E, and a transition function $T : E \to N \times N$. Furthermore, suppose there is also a cost function $c : E \to \mathbb{N}^+$ which returns a positive integer for the cost of traversing each edge.

(*i*) Which semiring lets us compute the length of the shortest path between any two nodes? [2 marks]

(*ii*) Augment the star-semiring to additionally compute the set of shortest paths between any two nodes. It is not necessary to prove that the semiring properties are satisfied. [5 marks]

(*iii*) How should the matrix whose star gives us the shortest paths be initialised? [3 marks]

3 Bioinformatics

(*a*) To assess sequence similarity, we compute the alignment between two DNA sequences.

(*i*) Compute the local alignment of the sequences GGTTATA, TATAGG with the following rules: match score = +4, mismatch = −3, gap penalty = −4. Discuss how the alignment depends on the value of match scores, mismatch and gap penalty. [5 marks]

(*ii*) Explain how DNA stores information in its natural biological role and how it could be used as a potential medium for artificial data storage. [4 marks]

(*b*) You conduct a self-experiment with two phases to investigate the effect of diet on gut microbiome diversity. In Phase 1, you follow a strict ketogenic diet (high fat, very low carbohydrate) for 3 weeks. In Phase 2, you switch to eating exclusively at McDonald's (high carbohydrate, processed food, low fibre) for 3 weeks. At the end of each phase, you collect a fecal sample and send it to a sequencing company for 16S rRNA sequencing. The company returns a list of bacterial species identified in each sample.

(*i*) Explain how you would compare bacterial composition between the two dietary phases. [5 marks]

(*ii*) Discuss the main limitations of this experimental design. [1 mark]

(*c*) Lloyd's algorithm (k-means) requires the number of clusters k to be specified in advance. Describe how you would choose an appropriate value of k. [5 marks]

4 Business Studies

You are a CEO of a software company. You have identified an AI tool that you believe can more effectively do the work currently done by lower performing test design engineers in some of your development teams. At this point, your options are to introduce this tool to some of your teams, or to use the AI tool to do the work of some of the engineers, or both.

(a) Taking into account employment law, outline your steps for reorganising teams. [4 marks]

(b) You decide that the AI tool should replace the work of a test engineer on a team. Choose a management framework to prepare for and implement this change in membership, and outline the steps according to the chosen framework. [8 marks]

(c) Suppose you instead decide to augment the team with the AI tool and change the roles of the team to adjust for this change. Choose a management framework different from Part (b) to prepare for and implement this change in membership, and outline the steps according to the chosen framework. [8 marks]

5 Cryptography

(*a*) How many different functions ($\rightarrow$) or permutations ($\leftrightarrow$) exist of each of the following types?

(*i*) $\{0,1\}^{256} \rightarrow \{0,1\}^{128}$

(*ii*) $\{0,1\}^{128} \leftrightarrow \{0,1\}^{128}$

(*iii*) $\mathbb{F}_{2^{64}} \rightarrow \mathbb{Z}_{10}^4$

(*iv*) $\mathbb{Z}_{pq} \leftrightarrow \mathbb{Z}_p \times \mathbb{Z}_q$

(*v*) $\mathbb{Z}_{13}^* \rightarrow \mathbb{Z}_{15}^*$

(*vi*) $\mathbb{E}(\mathbb{Z}_p, a, b) \rightarrow \{0,1\}$ [6 marks]

(*b*) X.509 certificates and Kerberos tickets are both trusted-third-party mechanisms for key distribution.

(*i*) Name six data fields found in a typical X.509 certificate. [3 marks]

(*ii*) Describe three key differences between X.509 certificates and Kerberos tickets. [3 marks]

(*iii*) What is the purpose of a ticket-granting ticket in Kerberos? [2 marks]

(*c*) Even though elliptic-curve group elements correspond to points in a two-dimensional space, they are often represented by three-dimensional coordinates in cryptographic implementations. Why is this done? [2 marks]

(*d*) Let (Gen, H_s) be a collision-resistant hash function. Is (Gen, H'_s) with $H'_s(x) = H_s(H_s(x))$ necessarily also collision resistant? Justify your answer. [4 marks]

6 Denotational Semantics

(*a*) Define *adequacy* and *full abstraction* for a semantics of PCF. [4 marks]

(*b*) Consider the flat domains $D = \mathbb{B}_\perp$ and $D = \mathbb{N}_\perp$. Show that the equality function $\mathrm{eq}_D \colon D \times D \to \mathbb{B}_\perp$ determined by

$$\mathrm{eq}_D \ x \ y = \begin{cases} \perp, & \text{if } x = \perp \text{ or } y = \perp \\ \text{true}, & \text{if } x \neq \perp,\ y \neq \perp \text{ and } x = y \\ \text{false}, & \text{if } x \neq \perp,\ y \neq \perp \text{ and } x \neq y \end{cases}$$

is definable in PCF by exhibiting terms $\mathsf{eq}_{\mathsf{bool}}$ and $\mathsf{eq}_{\mathsf{nat}}$ with denotation $[\![\mathsf{eq}_{\mathsf{bool}}]\!] = \mathrm{eq}_{\mathbb{B}_\perp}$ and $[\![\mathsf{eq}_{\mathsf{nat}}]\!] = \mathrm{eq}_{\mathbb{N}_\perp}$. Justify your answer. [5 marks]

(*c*) Consider the *parallel conditional* operations for $D = \mathbb{B}_\perp$ and for $D = \mathbb{N}_\perp$:

$$\mathrm{pif}_D \colon \mathbb{B}_\perp \times D \times D \to D$$

These are the unique continuous functions determined by the equations

$$\mathrm{pif}_D \ \text{true} \ x \ y = x \qquad \mathrm{pif}_D \ \text{false} \ x \ y = y \qquad \mathrm{pif}_D \ \perp \ x \ x = x$$

(*i*) What is the value of $\mathrm{pif}_D \perp x \ y$ for $x \neq y$? Justify your answer. [5 marks]

(*ii*) Prove that $\mathrm{pif}_{\mathbb{B}_\perp}$ is not definable in PCF. You may use the theorem that the *parallel or* operation $\mathrm{por} \colon \mathbb{B}_\perp \times \mathbb{B}_\perp \to \mathbb{B}_\perp$ is not definable. [3 marks]

(*iii*) Prove that $\mathrm{pif}_{\mathbb{N}_\perp}$ is not definable in PCF. [3 marks]

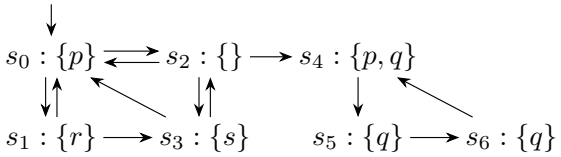
7 Hoare Logic and Model Checking

This question uses the following syntax for CTL:

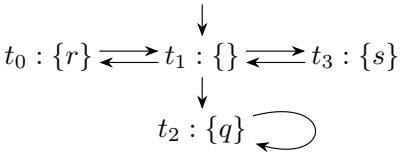
$$\psi \in \textbf{StateProp} ::= \bot \mid \top \mid \neg\psi \mid \psi_1 \wedge \psi_2 \mid \psi_1 \vee \psi_2 \mid \psi_1 \to \psi_2 \mid p \mid A\,\phi \mid E\,\phi$$
$$\phi \in \textbf{PathProp} ::= X\,\psi \mid F\,\psi \mid G\,\psi \mid \psi_1\,U\,\psi_2.$$

(*a*) Consider CTL formulae over $\mathsf{AP} = \{\, p \,\}$. A student writes down the formula $AF\,AX\,p$, intending to express the property that "p holds at some future state *strictly after* the current state". However, it actually specifies a different property. Explain why the formula doesn't match the intention, using a concrete temporal model, and specify the correct CTL formula. [5 marks]

(*b*) Consider a temporal model M over atomic propositions $\mathsf{AP} = \{\, p, q, r, s \,\}$, with states $S = \{\, s_0, s_1, s_2, s_3, s_4, s_5, s_6 \,\}$, initial state s_0, transitions and state labelling as shown in the diagram below on the left. For example, the diagram indicates that in state s_0 the atomic proposition p holds. Describe the meaning of CTL formula $AG\,(EF\,s \to (EX\,EX\,p))$, and explain whether it holds in M. [4 marks]



Temporal model M



Temporal model M'

(*c*) Consider now the right side of the figure, which shows a temporal model M'.

(*i*) Show that M' does not simulate M. [2 marks]

(*ii*) Make M' simulate M by adding as few transitions as possible. State the transitions you add, specify a simulation relation, and show that the modified M' simulates M. [4 marks]

(*d*) CTL formula φ and LTL formula Ψ, over the same atomic propositions, are equivalent if, for every temporal model M, $M \models \varphi$ if and only if $M \models \Psi$. Prove that there is no LTL formula equivalent to $AF(q \wedge AX\,p)$. Hint: in 1988, Clarke and Draghicescu showed the following fact, which you may find useful:

> Let φ be a CTL formula and Ψ be an LTL formula that is obtained by eliminating all the path quantifiers in φ; for example, $AF\,EX\,p$ becomes $F\,X\,p$ by eliminating the path quantifiers. There then exists an LTL formula equivalent to φ if and only if φ is equivalent to Ψ.

[5 marks]

8 Information Theory

(*a*) Compute the differential Entropy, $H(X)$, for $X \sim U[0.0, 2.0]$ and $X \sim U[0.0, 0.5]$. Provide an argument that differential entropy has meaning, based on the physical interpretation of these values. [4 marks]

(*b*) Consider the Gaussian channel $Y = X + Z$ where X is the input signal with power constraint $E[X^2] \leq P$, and Z is noise-independent of X, with $Z \sim \mathcal{N}(0, N)$.

(*i*) Express the Mutual Information $I(X;Y)$ of this channel in terms of the differential entropies $H(Y)$ and $H(Z)$. [2 marks]

(*ii*) Explain why achieving the capacity of this channel requires X to have a Gaussian distribution. You may assume without proof that the Gaussian distribution maximises differential entropy for a fixed variance. [3 marks]

(*iii*) Derive the Shannon-Hartley theorem for this channel. You may assume the capacity of the Gaussian channel is $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$. [3 marks]

(*c*) Consider sending information over two parallel, independent Gaussian channels with noise variances N_1 and N_2 respectively. The transmitter has a total power budget P, which must be distributed between the two channels ($P = P_1 + P_2$).

(*i*) Show that for an optimal solution where both channels are used, the following holds:

$$P_1 + N_1 = P_2 + N_2$$

You may assume the total capacity is the sum of the individual capacities: $C_{total} = \frac{1}{2}\log\_2\left(1 + \frac{P_1}{N_1}\right) + \frac{1}{2}\log\_2\left(1 + \frac{P_2}{N_2}\right)$. [4 marks]

(*ii*) A specific system has $N_1 = 2$ W and $N_2 = 10$ W and $P = 6$ W. Demonstrate that the optimal power allocation cannot be calculated using the relation derived in Part $(c)(i)$, explain why, and state the optimal power distribution. [4 marks]

9 Machine Learning and Bayesian Inference

Consider the following learning problem. Examples are pairs (x, c) with single, real-valued feature $x \in [0, 1]$ and class label $c \in \{0, 1\}$. A hypothesis h_g is specified by a parameter $g \in [0, 1]$ and has likelihood

$$\Pr(C = 1|h_g; x) = gx.$$

We fix two parameter values g_1 and g_2 and only consider the two corresponding hypotheses. The prior on the hypothesis is specified by a hyperparameter $p = \Pr(h_{g_1})$.

(*a*) Give at least two advantages and two disadvantages of the Bayesian approach to supervised machine learning. [5 marks]

(*b*) For a training set with examples $\mathbf{x}^T = [x_1, x_2, \ldots, x_m]$ and labels $\mathbf{c}^T = [c_1, c_2, \ldots, c_m]$, show that the Bayes prediction can be expressed as

$$\Pr(C = 1|\mathbf{c}; x, \mathbf{x}) = \frac{x}{Z}\left(g_1 p \Phi(h_{g_1}) + g_2(1 - p)\Phi(h_{g_2})\right)$$

where the function Φ and quantity Z should be defined in your answer. Start by stating the formula for the Bayes prediction $\Pr(C|\mathbf{c}; x, \mathbf{x})$ in this instance. [10 marks]

(*c*) Explain how the *evidence* can be employed as a means of estimating hyperparameters. Illustrate your answer by identifying the evidence in your answer to Part (*b*) and explaining how it can be used to estimate p when the latter is treated as a hyperparameter. [5 marks]

10 Optimising Compilers

(*a*) Explain the concept of *dominance*, including *dominance frontier*. [2 marks]

(*b*) You are given a RISC-V assembly code program **P**. Arguments are passed in registers `a0` and `a1`. The return value is returned in `a0`.

```
A: mv t0, a0       // t0 = a0 -- Move argument 0
   mv t1, a1       // t1 = a1 -- Move argument 1
   li t3, 0        // t3 = 0, Fallthrough
B: beq t3, t1, G   // Jump to G if `t3 == t1`, else fallthrough.
C: li t5, 2        // t5 = 2
   li t6, 0        // t6 = 0
   remu t4, t0, t5 // t4 = t0 % t5 (unsigned)
   bne t4, t6, E   // Jump to E if `t4 != t6` else fallthrough.
D: j F             // Jump to F
E: add t0, t0, t4  // Fallthrough
F: li t7, 2        // t7 = '2'
   add t3, t3, t7  // t3 = t3 + t7
   j B             // Jump to B.
G: mv a0, t0       // a0 = t0 -- Move result to return register.
   ret             // Return
```

(*i*) Draw the flow graph and dominator tree of **P**. [6 marks]

(*ii*) Assume **P** is run with input `a1 = 1`. How does **P** behave, and how does the execution behaviour of **P** impact the dominance property of **P**? [2 marks]

(*iii*) Turn the program into SSA form by renaming a minimal set of variables and placing a minimal set of PHI-nodes. To get multiple versions of the same register, rename `rX` to `rX0`, `rX1`, `rX2`, ... (e.g. `r0` to `r00`, `r01`, ...). For PHI-nodes, use the instruction format `phi rOUT, rINA (LABELA), rINB (LABELB)` (e.g. `phi r51, r50 (A), r51 (B)`). [6 marks]

(*c*) Compare the result size of a *reaching definitions* analysis (without storage optimisations) with the size it takes to encode reaching definitions via SSA, e.g. in numbers of use-def pairs. [2 marks]

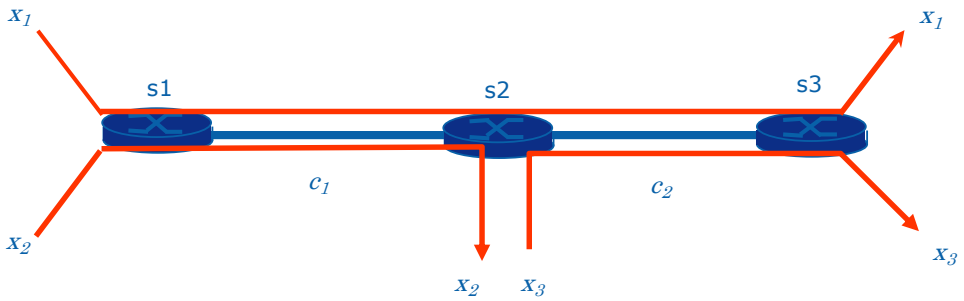
(*d*) Give a worst-case example program that demonstrates the size difference between reaching definitions with and without SSA, and explain what is responsible for this difference. Use C-style pseudocode with switch statements and a special function call `phi` that takes switch labels and variables and returns a variable. The `phi` function mirrors the semantics of a PHI-node in C. [2 marks]

11 Principles of Communications

(*a*) Open-loop flow control systems typically employ packet scheduling to ensure isolation and fairness. What is the key purpose of the associated component functions of admission control and policing? [5 marks]

(*b*) The figure shows a simple network with three switches s1, s2, and s3. The network has the following characteristics:

Links c_1, c_2 are both 1 Mbps, with 10ms one-way propagation delay.
Ingress ports at switches s1 and s2 are both 10 Mbps.
There are three flows x_1, x_2, and x_3 with the following characteristics:
for x_1 100 packets per second, 1 Kbyte packets (1K=1000), bursty;
for each of x_2 and x_3, 20 packets per second, 400 byte packets, constant rate.



(*i*) What is the worst case end-to-end delay for flow x_2 induced at s1 by flow x_1's burstiness? Assume that the routers implement a simple, per-flow Weighted Round Robin scheduler which is fair on a packet-rate basis. Clearly state any other assumptions you make in your answer. [8 marks]

(*ii*) Flow x_3 joins the network at switch s2 where x_2 leaves. Can we make one simple assumption to work out what latency x_3 will experience? [2 marks]

(*c*) Discuss the impact on worst case end-to-end latencies experienced by the three traffic flows if the routers implemented FIFO/FCFS, rather than Weighted Round Robin scheduling. [5 marks]

12 Quantum Computing

(a) Consider the following Hermitian matrix.

$$\mathrm{H} = \frac{\pi}{2\sqrt{2}} \begin{bmatrix} \sqrt{2}-1 & -1 \\ -1 & \sqrt{2}+1 \end{bmatrix}$$

(i) Define what it means for a matrix to be Hermitian. What property do the eigenvalues of Hermitian matrices have? [2 marks]

(ii) Verify that $\begin{bmatrix} 1 \\ \sqrt{2}-1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -\sqrt{2}-1 \end{bmatrix}$ are eigenvectors of H and find its eigenvalues. [4 marks]

(iii) Find the eigenvectors of the matrix $e^{-i\mathrm{H}t}$, where t is a positive real number. [1 mark]

(iv) What kind of matrix is $e^{-i\mathrm{H}t}$? [1 mark]

(b) Quantum states may be perfectly distinguished if they are orthogonal. Let $|\psi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$, $|\phi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $|\omega\rangle = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

(i) Show that $|\psi\rangle$ and $|\phi\rangle$ are orthogonal, and that $|\psi\rangle$ and $|\omega\rangle$ are not orthogonal. [3 marks]

(ii) Find a unitary matrix which, when applied to the states $|\psi\rangle$ and $|\phi\rangle$, allows them to be perfectly distinguished by measurement in the computational basis. [2 marks]

(iii) If one attempts to distinguish the non-orthogonal states $|\psi\rangle$ and $|\omega\rangle$, what is the probability of correctly inferring the state if the optimal strategy is used? [2 marks]

(iv) Give two measurement bases to distinguish $|\psi\rangle$ and $|\omega\rangle$ such that in each basis one measurement outcome has the property that the state is known with certainty. In your answer you should indicate the measurement outcome and the state for each measurement basis. [5 marks]

13 Types

(*a*) Give the type of the Church numerals in System F. [2 marks]

(*b*) Give an existential type that models the following OCaml module interface:

```
module type NatList = sig
  type t

  val nil : t
  val cons : nat -> t -> t
  val fold : t -> 'a -> (nat -> 'a -> 'a) -> 'a
end
```

[3 marks]

(*c*) Give a term with the existential type you defined in Part (*b*), as a term in System F extended with existential types and products. [5 marks]

(*d*) In the typed lambda calculus augmented with state but no monadic control, write a function $f : ((\mathbb{N} \to 1) \to 1) \to (\mathbb{N} \to 1) \to \mathbb{N}$. The function call $f\ k\ g$ should return n, where n is the number of times that g is invoked by k in the function call $k\ g$. [3 marks]

(*e*) A dependent type theory is both a system of formal logic and a programming language. Consider the following piece of Agda code, where `List A` is the type of lists with element type `A`.

```
X : (P : List A → Set) →
    (P []) →
    ((x : A) → (xs : List A) → P xs → P (x :: xs)) →
    (xs : List A) → P xs
X P base step [] = base
X P base step (x :: xs) = step x xs (X P base step xs)
```

(*i*) Explain what this function is, understood as a theorem of logic. [3 marks]

(*ii*) Explain what this function is, understood as a functional program. [4 marks]

END OF PAPER