

Question 1 [8 marks]

Below shows a pseudocode of insertion sort with blanks.

```
PROCEDURE InsertionSort(A : ARRAY[0 : N - 1] OF INTEGER)
    // using 0-indexing

    FOR i ← .....(1)..... TO .....(2).....
        temp ← A[i]
        j ← .....(3).....

        WHILE j >= 0 AND temp < .....(4).....
            .....(5)..... ← A[j]
            j ← .....(6).....
        ENDWHILE

        .....(7)..... ← temp
    NEXT i
ENDPROCEDURE
```

Write your answer for each blank **(1)** through **(7)** clearly. [3]

State the worst-case time complexity of insertion sort and briefly explain what that implies. [2]

Insertion sort is effective for small arrays due to its simplicity and low overhead. However, for larger arrays, more efficient algorithms such as merge sort are preferred. Describe how merge sort works and state its time complexity. [3]

Answers:

(a) (1) 1, (2) $N - 1$, (3) $i - 1$, (4) $A[j]$, (5) $A[j + 1]$, (6) $j - 1$, (7) $A[j + 1]$

(b) Worst-case time complexity: $O(n^2)$

For large datasets, the number of operations required for insertion sort grows quadratically with input size.

(c) Merge sort is a divide-and-conquer algorithm that sorts an array by recursively breaking it down and then combining them in sorted order.

- Divide the array in half repeatedly until each subarray contains only one element.

- Combine adjacent subarrays back together in sorted order. Continue this process until the entire sorted array is constructed.

Time complexity: $O(n \log n)$

Question 2 [15 marks]

(a) State what is meant by data integrity. [1]

Sonars Service Centre handles repairs of their proprietary Sonars audio products. When customers bring equipment for repair, technicians would need to enter the product serial number into the system. Each serial number consists of 10 digits, where the last digit is a check digit calculated using this algorithm:

1. In the first pass, multiply each of the first 9 digits of the serial number by weights: 9, 8, 7, 6, 5, 4, 3, 2, 1 respectively (e.g. position 1 gets weight 9). Sum these products to obtain subtotal.
2. In the second pass, using the same 9 digits of the serial number, multiply each digit by the sum of its own value and its position number (e.g. if the first digit is 5, we get a product of $5 \times (5 + 1) = 30$ for that digit). Sum these products to obtain another subtotal.
3. Add the two subtotals from the two passes to obtain a combined total.

4. Divide the combined total by 13 to obtain the remainder. Subtract the remainder from 13 to derive the check digit.

(b) (i) It is given that 284736197X is a valid serial number, where X represents the check digit. Calculate the value of X, showing your workings. [3]

(ii) Briefly explain why check digit alone is insufficient to ensure data integrity. [1]

(iii) Describe two data verification methods that can be used when manually entering serial numbers into the system. [2]

When a Sonars amplifier system arrives for diagnosis, the technician checks for these symptoms:

- whether the power LED is OFF
- whether the audio output is abnormal
- whether the amplifier is overheating

The following describes the actions the technician would take:

If the power LED is off, check the power supply and fuse only regardless of other symptoms. If the power LED is on and the audio output is abnormal (but not overheating), test the input connections. If the amplifier is overheating, clean the ventilation system and fans. If the audio is abnormal and the amplifier is overheating, also check the output transistors, as they may be faulty. If none of the 3 symptoms are present, perform exterior cleaning.

It is desired to draw a decision table for the above information.

(c)(i) Explain what a decision table is. [2]

(ii) Create a complete decision table based on the description above. [4]

(iii) Show a simplified decision table without redundancies. [2]

Answers:

(a) Data integrity refers to the accuracy, consistency and reliability of data throughout its entire lifecycle (creation, storage, transmission).

(b)(i) For the first pass, subtotal = $2 \times 9 + 8 \times 8 + 4 \times 7 + 7 \times 6 + 3 \times 5 + 6 \times 4 + 1 \times 3 + 9 \times 2 + 7 \times 1 = 219$

For the second pass, subtotal = $2 \times (2+1) + 8 \times (8+2) + 4 \times (4+3) + 7 \times (7+4) + 3 \times (3+5) + 6 \times (6+6) + 1 \times (1+7) + 9 \times (9+8) + 7 \times (7+9) = 560$

Combined total = $219 + 560 = 779$

Remainder = $779 \text{ MOD } 13 = 12$

Check digit, $X = 13 - 12 = 1$

(ii) Check digit cannot catch all possible data entry errors, such as multiple errors that might cancel each other out. While it verifies mathematical correctness based on the algorithm, it cannot detect whether the serial number exists for an actual product.

(iii) Double entry: requiring user to enter the same serial number twice in separate input fields. The system compares both entries to ensure they are identical before accepting the data.

Visual confirmation: displaying a message box showing the entered number and prompting the user to confirm that it matches the serial number on the physical device.

(c)(i) A decision table is a tabular method to document decision-making logic. It systematically shows all possible combinations of conditions, with each of the corresponding actions to take.

(ii) Decision table with 8 rules:

Rule	1	2	3	4	5	6	7	8
Conditions								
Power LED is OFF	Y	Y	Y	Y	N	N	N	N
Audio output is abnormal	Y	Y	N	N	Y	Y	N	N
Amplifier is overheating	Y	N	Y	N	Y	N	Y	N
Actions								
Check power supply and fuse	X	X	X	X				
Test input connections						X		
Clean ventilation and fans					X		X	
Check output transistors					X			
Perform exterior cleaning								X

(iii) Simplified decision table:

Conditions					
Power LED is OFF	Y	N	N	N	N
Audio output is abnormal	---	Y	Y	N	N
Amplifier is overheating	---	Y	N	Y	N
Actions					
Check power supply and fuse	X				
Test input connections			X		
Clean ventilation and fans		X		X	
Check output transistors		X			
Perform exterior cleaning					X

Question 3 [12 marks]

Harley Thorman Corporation is developing an inventory management system for its electronic products using Object-Oriented Programming (OOP).

The main types of electronic devices that the corporation sells are display devices, audio devices and computing devices.

Display devices include televisions, computer monitors and screen projectors. Audio devices include speaker systems and headphones. Computing devices include desktops, laptops and tablets.

(a) Draw a class diagram for this system showing classes and their inheritance relationships. Do **not** include any attributes or methods in the diagram. [3]

(b) Explain why inheritance is an important feature in OOP. [1]

(c) Suggest which class each of the following attributes should be placed in, explaining your choice:

- CPU model
- product serial number. [3]

(d) Suggest a possible private method that one might include in the OOP system, explaining why keeping it private would be reasonable. [2]

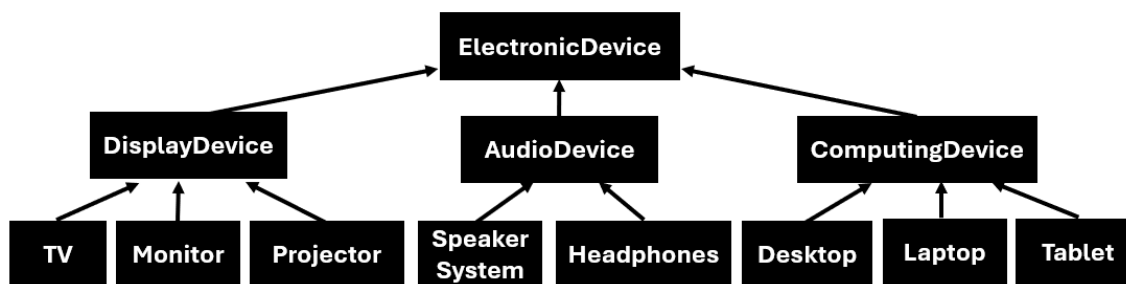
There are different types of TV display technologies, each with distinct features. LCD TVs use liquid crystal displays with backlighting, while LED TVs use LED backlighting for better energy efficiency. OLED TVs use organic light-emitting diodes that allow pixels to turn completely off, providing superior contrast ratios. QLED TVs use quantum dot technology to enhance brightness and colour accuracy. What distinguishes them is the panel technology type, which results in differences in display characteristics such as contrast ratio, peak brightness and colour gamut coverage.

(e) Suggest how the OOP system should be updated to accommodate these different TV display technologies, justifying the approach chosen. [2]

(f) Explain what is meant by polymorphism in OOP. [1]

Answers:

(a) Class diagram:



(b) Inheritance enables developers to reuse code efficiently, as common attributes and methods can be defined in a parent class, and all subclasses would inherit these functionalities. It eliminates the need to duplicate code across multiple classes. Updates made to the parent class automatically propagate to all subclasses. Therefore, it makes the codebase more maintainable and consistent.

(c) CPU model should be placed in the **ComputingDevice** class, as typically only computing devices (desktops, laptops, tablets) would include CPU model information in the specifications.

Product serial number should be placed in the **ElectronicDevice** base class, as all electronic products need unique serial numbers for identification.

(d) `validateSerialNumber()` would be a possible private method to include in the **ElectronicDevice** class. It should be private because it is an internal validation process that other classes don't need to access directly.

(e) We should add an attribute 'panel_technology' (LCD/LED/OLED/QLED) to the **Television** class, rather than creating separate subclasses for each TV type. They would have largely similar methods and attributes, differing only in specifications.

(f) Polymorphism refers to the ability of a method to exhibit different behaviours for different classes within the class hierarchy. This is achieved through subclasses overriding methods inherited from their parent class.

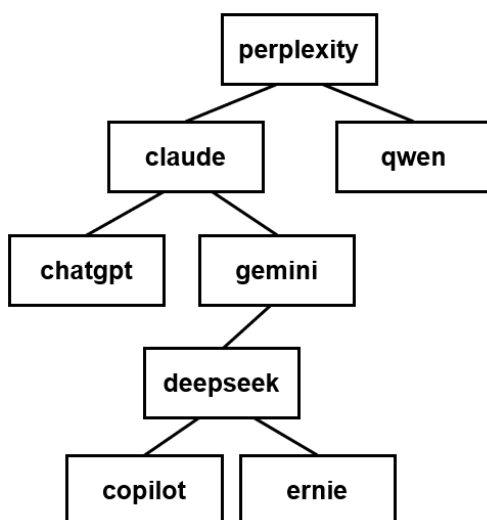
Question 4 [16 marks]

You are given the following sequence of words (which represent different popular generative AI platforms): perplexity, claude, gemini, deepseek, chatgpt, copilot, qwen, ernie.

- (a) These words are inserted one by one into an empty binary search tree (BST) in the given order, with nodes arranged based on lexicographical (alphabetical) order. Draw the complete BST after insertion. [2]
- (b) List the nodes visited in **pre-order** and **post-order** traversal. [2]
- (c) What property does the **in-order** traversal have for the BST? [1]
- (d) If we replace the word 'perplexity' with 'grok' in the insertion sequence while maintaining the same order of insertions, will the resulting BST structure be different? Explain your reasoning without reconstructing the BST. [2]
- (e) Write, in pseudocode, a function `Search(root, target)` to perform a binary tree search for the value `target`. It should return a boolean indicating whether the target value is present in the tree. [4]
- (f) Write, in pseudocode, a recursive function `CountNodes(root)` that returns the total number of nodes present in the BST. [3]
- (g) For the case where many words need to be stored, discuss the advantages of using a binary search tree compared to a sorted linked list. [2]

Answers:

(a)



(b) **pre-order**: perplexity, claude, chatgpt, gemini, deepseek, copilot, ernie, qwen

post-order: chatgpt, copilot, ernie, deepseek, gemini, claude, qwen, perplexity

(c) in-order traversal of a BST produces nodes in sorted alphabetical order.

(d) No, the BST structure will not be different. Replacing 'perplexity' with 'grok' maintains the same relative lexicographical ordering with respect to other words. All words on the left subtree of 'perplexity' are still lexicographically smaller than 'grok', and all words on the right subtree are larger than 'grok'. Therefore, no changes in the final tree structure will result.

(e)

```
FUNCTION Search(root, target) RETURNS BOOLEAN
    // Base case (empty tree or reached end)
    IF root = NULL THEN
        RETURN FALSE
    ENDIF

    // Found target
```

```

IF root.data = target THEN
    RETURN TRUE
ENDIF

IF target < root.data THEN
    // search left
    RETURN Search(root.left, target)
ELSE
    // search right
    RETURN Search(root.right, target)
ENDIF
ENDFUNCTION

```

(f)

```

FUNCTION CountNodes(root) RETURNS INTEGER
    // Base case (empty tree)
    IF root = NULL THEN
        RETURN 0
    ENDIF

    // Count nodes in subtrees
    leftCount ← CountNodes(root.left)
    rightCount ← CountNodes(root.right)

    RETURN 1 + leftCount + rightCount
ENDFUNCTION

```

(g)

Significantly faster search (or insertion) operations: For BST, time complexity is $O(\log n)$ on average for word lookup. This is because through one comparison at a node, you can eliminate half the remaining search space.

For a sorted linked list, it requires $O(n)$ time complexity as it requires sequential traversal starting from the head. This difference is significant for large word collections.

Question 5 [13 marks]

Refer to the following recursive function P:

```

01 FUNCTION P(x : REAL, a : INTEGER) RETURNS REAL
02     IF a = 0 or a = 1 THEN
03         RETURN x
04     ENDIF
05
06     IF a MOD 2 = 0 THEN
07         r ← P(x, a / 2)
08         RETURN r * r
09     ELSE
10         RETURN x * P(x, a - 1)
11     ENDIF
12 ENDFUNCTION

```

(a) With reference to relevant line numbers in the pseudocode, explain the three features seen here that make P a recursive function. [3]

The author of the function intends for it to work for all non-negative integer values of a.

(b) Identify the intended purpose of this function. [2]

(c) Upon testing, the author realises that there is a logic error in the code. Identify this error and state how it should be corrected. [2]

(d) Draw a recursion trace diagram to illustrate the recursive steps and the return value calculations when computing $P(3.0, 7)$. [3]

(e) State what would happen if $P(3.0, -1)$ is executed. [1]

(f) Write down what you would add to the existing pseudocode for the function to handle negative integer values of a correctly. [2]

Answers:

(a) The base case is seen in line 02. A terminating condition is provided that stops further function calls.

Lines 07 and 10 show the recursive call, where the function calls itself.

Lines 07 and 10 also show function calls that move closer to the base case.

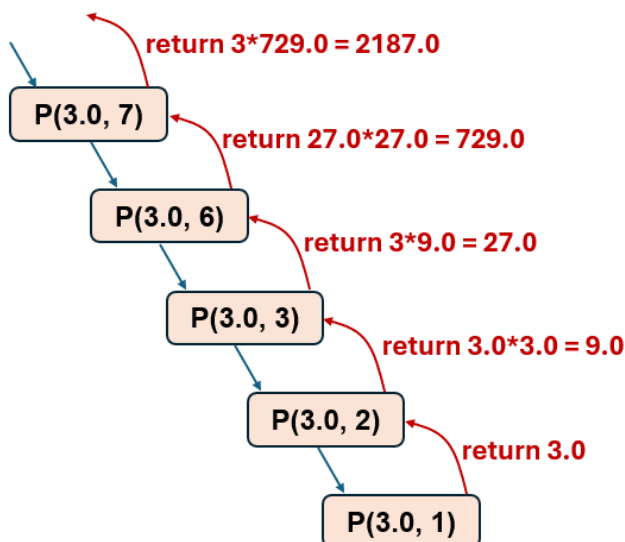
(b) The function is intended to calculate x raised to the power of the value a .

(c) When $a = 0$, the function wrongly returns x instead of 1.

Lines 2 – 3 should be changed to:

```
IF a = 0 THEN
    RETURN 1
ENDIF
IF a = 1 THEN
    RETURN x
ENDIF
```

(d) Recursion trace:



(e) It calls $P(3.0, -2)$, which then calls $P(3.0, -1)$ again. This cycle repeats, leading to a stack overflow error, a type of runtime error.

(f) Before line 02, add the following code:

```
IF a < 0 THEN
    RETURN 1.0 / P(x, -1*a)
ENDIF
```

Question 6 [19 marks]

Yee Kee Yah Ltd operates 3 furniture outlets in Singapore. The company has decided to implement a computerised inventory management system with a web app interface.

(a) The IT team has recommended using client-server architecture for each outlet's inventory system.

(i) Explain what is meant by a client-server model in this context. [2]

(ii) Identify three types of software required to implement this client-server system. [3]

(iii) Write down one advantage of adopting a client-server system in this context. [1]

(b) Currently, each outlet operates independently. To improve coordination, the furniture company wants to connect all three outlets to its headquarters so they can share inventory data and coordinate stock transfers.

(i) Explain how LAN and WAN technologies would work together to support the company's inventory management operations and describe the role of routers in this network setup. [3]

(ii) Explain how an intranet would benefit the operations. [2]

(iii) State two methods for ensuring the security of access to the inventory management application. [2]

(iv) Suggest two methods to secure the network infrastructure. [2]

(c) The company also keeps track of customer purchase history, delivery addresses and payment information.

(i) Explain the difference between data backup and data archiving of the customer and sales data. [2]

(ii) Describe two specific PDPA compliance measures Yee Kee Yah Ltd must implement when handling customer data. [2]

Answers:

(a)(i) A client-server model is a network architecture where the server stores and manages the inventory database, while clients (individual computers at workstations) request data and services from the server. The server processes client requests and sends back the requested data.

(ii)

Database management system (DBMS) software – to manage the inventory database

Web server software – to serve the web app and handle client requests

Web browser software – for clients to access the inventory system

(iii) Centralised data storage ensures all staff access the same up-to-date inventory information.

(b)(i) Each outlet and the headquarters uses a LAN to connect devices within that location. A WAN connects all four LANs (three outlets plus HQ) together across these distant locations. Routers connect each location's LAN to the WAN and direct data packets between different networks, enabling efficient inventory data sharing and stock transfer coordination between all outlets and headquarters.

(ii) An intranet provides a private internal network accessible only to company staff across all locations, allowing secure sharing of inventory data and internal communications. External parties cannot access the intranet, which helps protect sensitive company information and inventory data from external threats while enabling coordinated operations.

(iii) Implement user authentication (through usernames and passwords, 2FA) to ensure only staff can access the application.

Use role-based access control (setting different permission levels for different staff roles), reducing the risk of unauthorised data access and changes.

(iv) Deploy firewall to monitor and filter incoming and outgoing network traffic.

Use data encryption (such as using VPN) to secure data transmitted across the WAN, preventing interception.

(c) (i) Data backup is the process of making a copy of the current, active customer and sales data so it can be quickly restored in case the original data is lost or corrupted. It is used for disaster recovery.

Data archiving involves moving older, inactive data to a separate storage location for long-term retention and compliance purposes. It is meant for future reference.

(ii) Any two reasonable answers with explanations. For example:

Consent obligation: The company must inform customers about the purpose of collecting their personal data and obtain their consent before collecting, using or disclosing it.

Data protection obligation: The company must secure customer data by using safeguards such as encryption and regular security audits

Access and correction obligation: The company should provide customers with the right to access their personal data and request corrections if the data is inaccurate.

Question 7 [15 marks]

A university has several departments, each with multiple lecturers. A record of courses taught by each lecturer has been set up using a relational database. Each course is only taught by one lecturer.

Students can enrol in multiple courses. They may also re-enrol in the same course in another semester, particularly when they need to repeat the course due to unsatisfactory performance.

The following tables hold the data.

STUDENT (StudentID, StudentName, StudentEmail, StudentPhone)

DEPARTMENT (DeptID, DeptName, DeptAddress)

COURSE (CourseID, CourseName, LecturerID, MaxCapacity)

LECTURER (LecturerID, LecturerName, DeptID)

ENROLMENT (StudentID, CourseID, EnrolmentDate, Grade, Status)

(a) Copy the above table definitions, notating all primary keys and foreign keys. [4]

(b) Draw an Entity-Relationship (E-R) diagram to represent the above. [3]

(c) There is an address field in this database. Explain why storing the address as a single field is not good database design. [1]

A report is to be generated showing the courses taken by a student and the corresponding grades.

Student: Michael Tan

Course ID	Course Name	Grade
CS1231	Discrete Structures	A-
DSA5101	Introduction to Big Data for Industry	B+
DSA5102	Foundations of Machine Learning	A
DSA5105	Principles of Machine Learning	A-
DSA5204	Deep Learning and Applications	B
DSA5206	Advanced Topics in Data Science	B+

(d) Write an SQL query that would enable us to extract the required information above for the student Michael Tan, sequenced in ascending order by course ID. [3]

The university has decided to email student grade reports as PDF documents with digital signatures.

(e)(i) Explain the purpose of a digital signature. [2]

(ii) Describe the process of digitally signing a PDF document by the sender, and the process conducted by the receiver. [4]

Answers:

(a)

STUDENT (StudentID, StudentName, StudentEmail, StudentPhone)

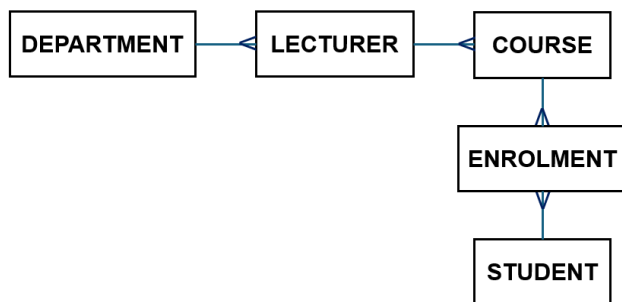
DEPARTMENT (DeptID, DeptName, DeptAddress)

COURSE (CourseID, CourseName, LecturerID, MaxCapacity)

LECTURER (LecturerID, LecturerName, DeptID)

ENROLMENT (StudentID, CourseID, EnrolmentDate, Grade, Status)

(b) E-R diagram:



(c) It violates the principle of atomicity required in 1NF. Address contains multiple separate components (e.g. block number, street, etc.) that should be stored in individual fields to allow for efficient searching and manipulation.

(d) SQL query:

```
SELECT COURSE.CourseID, COURSE.CourseName, ENROLMENT.Grade
FROM STUDENT
INNER JOIN ENROLMENT ON STUDENT.StudentID = ENROLMENT.StudentID
INNER JOIN COURSE ON ENROLMENT.CourseID = COURSE.CourseID
WHERE STUDENT.StudentName = 'Michael Tan'
ORDER BY COURSE.CourseID;
```

Alternatively:

```
SELECT COURSE.CourseID, COURSE.CourseName, ENROLMENT.Grade
FROM STUDENT, ENROLMENT, COURSE
WHERE STUDENT.StudentID = ENROLMENT.StudentID
      AND ENROLMENT.CourseID = COURSE.CourseID
      AND STUDENT.StudentName = 'Michael Tan'
ORDER BY COURSE.CourseID;
```

(e) (i) It ensures:

Authentication (confirming identity of the sender)

Data integrity (verifies the document has not been tampered with)

Non-repudiation (prevents the sender from denying that he signed the document)

(ii) Sender:

1. The PDF document is passed through a cryptographic hash function to create a unique hash value.
2. The hash is encrypted using the sender's private key to create the digital signature, which is attached to the PDF document.

Receiver:

1. The receiver extracts the digital signature and decrypts it using the sender's public key to obtain the original hash.
2. He also generates a hash of the received document using the same hash function and compares it with the decrypted hash. If they match, the digital signature is valid.