115學年度 學士後醫學系招生考試

計算機概論與程式設計試題封面

考試開始鈴響前，請勿翻閱本試題！

★考試開始鈴響前，請注意：

一、除准考證、應考文具及一般手錶外；行動電話、穿戴式裝置及其他物品均須放在臨時置物區。

二、請務必確認行動電話已取出電池或關機，行動電話及手錶的鬧鈴功能必須關閉。

三、就座後，不可擅自離開座位或與其他考生交談。

四、坐定後，雙手離開桌面，確認座位號碼、答案卡號碼與准考證號碼相同，以及抽屜中、桌椅下或座位旁均無非考試必需用品。如有任何問題，請立即舉手反應。

五、考試開始鈴響前，不得翻閱試題本或作答。

六、考試全程不得吃東西、喝水及嚼食口香糖。

七、違反上述規定，依「筆試規則及違規處理辦法」議處。

★作答說明：

一、考試時間：100 分鐘。

二、本試題(含封面)共 **12** 頁，如有缺頁或毀損，應立即舉手請監試人員補發。

三、本試題單選題共 30 題、申論題 4 題，共計 100 分；每題單選題答錯倒扣，不作答不計分。

四、單選題答題依題號順序劃記在答案卡上，寫在試題本上無效；答案卡限用 2B 鉛筆劃記，若未按規定劃記，致電腦無法讀取者，考生自行負責。

五、申論題部分以「答案卷」作答，作答時不得使用鉛筆，違者該科答案卷不予計分；限用黑色或藍色墨水的筆書寫。

六、試題本必須與答案卡一併繳回，不得攜出試場。

115 學年度學士後醫學系招生考試
計算機概論與程式設計試題

【單選題】每題 2 分，共計 60 分。答錯 1 題倒扣 0.5 分，倒扣至本大題零分為止，未作答，不給分亦不扣分。

1. In a direct-mapped cache system with 16 cache lines, a 32-bit memory address is divided into Tag, Index, and Offset fields. Assuming a byte-addressable memory and a block size of 64 bytes, how many bits are used for the Tag field?
   (A) 16 bits
   (B) 20 bits
   (C) 22 bits
   (D) 24 bits
   (E) 26 bits

2. A system implements the Banker's Algorithm for deadlock avoidance. There are 3 resource types (A, B, C). The current Available resources are (2, 1, 2). The current Allocation and Maximum matrices are as follows:

Process	Allocation (A, B, C)	Maximum (A, B, C)
P0	2, 0, 1	4, 2, 2
P1	3, 1, 2	5, 3, 3
P2	1, 2, 0	3, 3, 2

   Which of the following resource requests can be safely granted?
   (A) P0 requests (1, 1, 0)
   (B) P1 requests (2, 2, 1)
   (C) P2 requests (2, 1, 2)
   (D) P0 requests (2, 2, 1)
   (E) P1 requests (3, 1, 1)

3. Which statement best describes a key difference between how CNNs and Transformers extract features from input data?
   (A) CNNs rely on global attention mechanisms, while Transformers use fixed-size convolutional filters.
   (B) CNNs and Transformers extract features in identical ways but differ only in training speed.
   (C) Both CNNs and Transformers require manually designed filters for feature extraction.
   (D) CNNs capture local spatial patterns using convolution, while Transformers model global relationships using self-attention.
   (E) Transformers can only process sequential text data, while CNNs can only process images.

4. In Internet of Things (IoT) systems, which characteristic best describes the MQTT protocol?
   (A) It is a high-bandwidth protocol designed for video streaming.
   (B) It is mainly used for peer-to-peer file sharing.
   (C) It provides reliable communication by using a heavy connection setup.
   (D) It is optimized for low-bandwidth, low-power communication using a publish/subscribe model.
   (E) It replaces TCP/IP as the underlying network protocol.

5. About IPv4 header, which of the following is **INCORRECT**?
   (A) IPv4 header is 4 bytes alignment.
   (B) Differentiated Services Code Point (DSCP) is used for specifying differentiated services.
   (C) The IPv4 TTL is an upper bound of routers that an IP datagram can pass.
   (D) When receiving IPv4 fragments, the receiver can use header checksum field to reassemble the original packet.
   (E) IPv4 uses a 32-bit address space, which provides 4,294,967,296 (2^{32}) unique addresses.

6. Which of the following statements describing a time-complexity difference between a singly linked list and an array is **CORRECT**?
   (A) Deleting an element from the middle of a linked list (when the node is already known) takes O(1), while deleting from the middle of an array takes O(n).
   (B) Inserting an element at the beginning of a linked list takes O(n), while inserting at the beginning of an array takes O(1).
   (C) Accessing the k-th element in a linked list takes O(1), while accessing the k-th element in an array takes O(n).
   (D) Searching for a value in either a linked list or an array always takes O(1).
   (E) Appending an element to the end of a linked list without a tail pointer always takes O(1).

7. Which of the following best describes the concept of k-anonymity in the context of data privacy protection?
   (A) Encrypting sensitive data using strong cryptographic algorithms
   (B) Ensuring each record is indistinguishable from at least k-1 other records based on certain identifying attributes
   (C) Adding k random noise to query results to limit the influence of any single individual
   (D) Restricting access to data based on k user roles and permissions
   (E) Replacing personal identifiers with k random tokens to prevent identity disclosure

8. In data mining, consider a transactional database with 1,000,000 transactions. An itemset X appears in 12,000 transactions, and an itemset Y appears in 15,000 transactions.The combined itemset X∪Y appears in 9,000 transactions.Which of the following statements is **CORRECT**?
   (A) The confidence of the rule X→Y is 0.75.
   (B) The support of the rule X→Y is 0.75.
   (C) The lift of the rule X→Y is less than 1.
   (D) The rule X→Y is independent because its confidence equals the support of Y.
   (E) The conviction of the rule X→Y must be less than 1.

9. A processor uses a pipelined architecture with multiple stages. Which situation most directly causes a control hazard?
   (A) Two instructions require the same hardware resource at the same time.
   (B) An instruction depends on the result of a previous instruction.
   (C) Data must be transferred between cache and main memory.
   (D) The processor encounters a branch instruction whose outcome is not yet known.
   (E) An instruction has a longer execution time than others.

10. In an operating system, various synchronization mechanisms are used to coordinate concurrent processes. Which of the following statements about mutex locks and semaphores is **CORRECT**?
   (A) A mutex allows multiple processes to enter a critical section simultaneously as long as they release the lock quickly.
   (B) A counting semaphore can allow a limited number of processes to access a shared resource concurrently, based on its initial value.
   (C) A binary semaphore automatically prevents deadlock by detecting circular wait conditions.
   (D) A mutex lock increases system throughput by eliminating the need for context switching.
   (E) Semaphores guarantee fairness by ensuring processes access the critical section in strict FIFO order.

11. Standard Recurrent Neural Networks (RNNs) often suffer from the vanishing gradient problem. Which of the following architectural modifications is most effective at addressing this issue?
   (A) Increasing the learning rate
   (B) Adding more hidden layers
   (C) Using ReLU activation functions exclusively
   (D) Reducing the sequence length
   (E) Using LSTM (Long Short-Term Memory) units

12. What is the primary function of Network Function Virtualization (NFV) in modern networking architecture?
    (A) It centralizes routing decisions into a single hardware controller.
    (B) It virtualizes network services that traditionally run on dedicated hardware devices.
    (C) It increases physical bandwidth by upgrading transmission media.
    (D) It replaces packet-switching mechanisms with circuit-switching mechanisms.
    (E) It encrypts all network traffic at the data link layer by default.

13. Consider the following adjacency matrix representing a weighted graph:

	0	1	2	3
0	0	4	∞	2
1	4	0	3	∞
2	∞	3	0	1
3	2	∞	1	0

    Using Dijkstra's algorithm to find the shortest path from node 0 to node 2, what is the shortest distance?
    (A) 2
    (B) 3
    (C) 4
    (D) 5
    (E) 7

14. You are given a Python function that is supposed to calculate the sum of all integers from 1 to n using recursion. However, there is a mistake in the code. Identify the mistake and provide the corrected version of the code.

```
def sum_to_n(n):
    if n == 1:    # base case
        return 1
    else:
        return n + sum_to_n(n)
```

    (A) Change sum_to_n(n) to sum_to_n(n-1) in the recursive call
    (B) Change return 1 to return 0 in the base case
    (C) Change n + sum_to_n(n) to n * sum_to_n(n) in the recursive call
    (D) Change the base case from n == 1 to n == 0
    (E) Change return 1 to return n in the base case

15. Which statement best explains why hashing is not suitable for encrypting data that needs to be recovered later?
    (A) Hash functions are computationally expensive to compute.
    (B) Hash functions generate variable-length outputs.
    (C) Hash functions are only applicable to textual data.
    (D) Hash functions require a shared secret key.
    (E) Hash functions are one-way functions and cannot be reversed.

16. Which of the following statements about major data mining algorithms is **CORRECT**?
    (A) The Apriori algorithm improves efficiency by generating candidate itemsets using a depth-first search strategy and eliminating candidates based on information gain.
    (B) The k-means clustering guarantees convergence to the global optimum solution because it minimizes within-cluster variance at each iteration.
    (C) Decision tree algorithms such as ID3 and C4.5 select attributes using impurity measures (e.g., information gain or gain ratio) and recursively partition the data space.
    (D) DBSCAN requires the number of clusters k to be specified in advance and assumes clusters are spherical in shape.
    (E) Principal Component Analysis is a supervised dimensionality reduction technique that maximizes class separability.

17. Which of the following has the same in-order traversal sequence as its post-order traversal sequence?
    (A) Left skewed binary tree
    (B) Right skewed binary tree
    (C) Complete binary tree
    (D) Full binary tree
    (E) Perfect binary tree

18. Which of the following best explains the concept of a page fault in a virtual memory system?
    (A) A page fault occurs when a page is modified but not written back to disk.
    (B) A page fault occurs when a process tries to access a page that is not currently in physical memory.
    (C) A page fault occurs when the page table becomes full.
    (D) A page fault occurs when two processes request the same page simultaneously.
    (E) A page fault occurs when the operating system terminates a process.

19. Which of the following best describes the role of edge computing in AI-powered systems?
    (A) Processing all AI workloads only in large centralized cloud data centers
    (B) Improving AI accuracy by training all models exclusively on historical datasets
    (C) Reducing storage costs by deleting raw sensor data before any processing occurs
    (D) Using only wireless networks to transmit AI inference results
    (E) Processing data closer to where it is generated to reduce latency and improve real-time AI performance

20. Which of the following statements about TCP Reno is **INCORRECT**?
    (A) TCP Reno uses Slow Start to exponentially increase the congestion window at the beginning of a connection.
    (B) TCP Reno switches to Congestion Avoidance once the congestion window (cwnd) reaches the slow-start threshold (ssthresh).
    (C) TCP Reno enters Fast Recovery after receiving three duplicate ACKs.
    (D) TCP Reno reduces the congestion window (cwnd) to one segment after Fast Recovery.
    (E) Packet loss detected by timeout causes TCP Reno to enter Slow Start.

21. In a classic 5-stage CPU pipeline (IF, ID, EX, MEM, WB), what is the main effect of pipelining on program execution?
    (A) It reduces the latency of each individual instruction to less than one clock cycle.
    (B) It increases the CPU clock frequency without any need for hazard handling.
    (C) It improves overall throughput by allowing multiple instructions to be in different stages at the same time, though hazards may cause stalls or require forwarding.
    (D) It guarantees that the program always finishes in exactly $n + 4$ cycles for n instructions, regardless of branch instructions.
    (E) It eliminates the need for cache memory because instructions are executed continuously.

22. The arithmetic of Galois Field GF(2^8) can be implemented using polynomial multiplication modulo an irreducible polynomial. Assume the irreducible polynomial is $m(x)=x^8+x+1$. Compute the multiplication of the two byte values: 0x13 and 0x04 in GF(2^8).
    (A) 0x26
    (B) 0x4C
    (C) 0x98
    (D) 0x2F
    (E) 0x14

23. In supervised learning, what is the primary purpose of using a validation dataset?
    (A) To increase the size of the training data.
    (B) To evaluate the final performance of the model.
    (C) To tune model hyperparameters and detect overfitting.
    (D) To generate new features automatically.
    (E) To replace the test dataset.

24. Consider the following relation: Enrollment(StudentID, StudentName, CourseID, CourseName, InstructorName). Assume 1) Each student has only one StudentName; 2) Each course has only one CourseName; 3) Each course is taught by only one InstructorName; 4) A student can enroll in multiple courses.Which of the following decompositions correctly converts the relation into Third Normal Form (3NF)?
    (A) Enrollment(StudentID, CourseID)
    (B) Enrollment(StudentID, StudentName, CourseID), Course(CourseID, InstructorName)
    (C) Student(StudentID, StudentName, CourseID),
    Course(CourseID, CourseName, InstructorName)
    (D) Student(StudentID, StudentName), Course(CourseID, CourseName),
    Instructor(InstructorName, CourseID), Enrollment(StudentID, CourseID, InstructorName)
    (E) Student(StudentID, StudentName), Course(CourseID, CourseName, InstructorName),
    Enrollment(StudentID, CourseID)

25. Consider the following processes with their arrival times and CPU burst times (in milliseconds):

Process	Arrival Time	Burst Time
P1	0	6
P2	2	8
P3	4	7
P4	5	3

    Using non-preemptive Shortest Job First (SJF) scheduling, what is the Average Waiting Time (AWT)?
    (A) 5 ms
    (B) 5.5 ms
    (C) 6 ms
    (D) 6.5 ms
    (E) 7 ms

26. Natural Language Processing (NLP) is one of the most popular AI applications. What is the purpose of NLP tokenization?
    (A) To store text in a compressed binary format.
    (B) To translate text into another language.
    (C) To ensure that all sentences have the same length.
    (D) To remove all punctuation from text.
    (E) To convert raw text into smaller units such as words or subword tokens for model processing.

27. You are given a Python function that is supposed to implement the merge sort algorithm to sort a list in ascending order. However, there is a mistake in the code. Identify the mistake and provide the corrected version of the code.

```
def merge_sort(arr):
        if len(arr) <= 1:
            return arr
        mid = len(arr)
        left = merge_sort(arr[:mid])
        right = merge_sort(arr[mid:])
        return merge(left, right)
def merge(left, right):
        result = []
        i = j = 0
        while i < len(left) and j < len(right):
            if left[i] < right[j]:
                result.append(left[i])
                i += 1
            else:
                result.append(right[j])
                j += 1
        result.extend(left[i:])
        result.extend(right[j:])
        return result
```

Which of the following options correctly fixes the error?
    (A) Change result.append(left[i]) to result.append(right[j])
    (B) Change result.append(right[j]) to result.append(left [i])
    (C) Change while i < len(left) and j < len(right) to while i <= len(left) and j <= len(right)
    (D) Remove the recursive calls in merge_sort
    (E) Change mid = len(arr) to mid = len(arr) // 2

28. Which of the following statements is **CORRECT** regarding NP-complete problems?
    (A) If a polynomial-time algorithm is found for any NP-complete problem, then all problems in NP can be solved in polynomial time.
    (B) Every NP-complete problem has a polynomial-time algorithm that approximates the optimal solution within a constant factor.
    (C) NP-complete problems are those for which no polynomial-time verification algorithm exists.
    (D) A problem is NP-complete if and only if it can be reduced in polynomial time to every problem in NP.
    (E) If a problem is NP-complete, then it cannot be an NP-Hard problem.

29. Please calculate the value of 2^{2026} mod 13. Which of the following is **CORRECT**?
    (A) 0
    (B) 1
    (C) 7
    (D) 10
    (E) 11

30. In a decision tree for classification, which measure is commonly used to determine the best attribute to split on at each node?
    (A) Euclidean Distance
    (B) Cosine Similarity
    (C) Information Gain (Entropy Reduction)
    (D) Manhattan Distance
    (E) Pearson Correlation

【申論題】每題 10 分，共計 40 分。未作答或作答錯誤，不給分亦不扣分。

1. A medical research team is building a classifier to predict disease presence (Positive/Negative) based on patient data. After testing on 200 patients, they obtain the following confusion matrix:

		Predicted	
		Positive	Negative
Actual	Positive	72	18
	Negative	12	98

(1) (5%) Calculate the following metrics and explain what each represents (Show the result with percentage, e.g. 0.78256 should be stated as 78.26 %):

a. Accuracy

b. Precision

c. Recall (Sensitivity)

d. F1-Score

e. Specificity

(2) (5%) The team wants to adjust the classification threshold to reduce false negatives (missed diseases), even if it increases false positives. Explain which metric should be prioritized and why this trade-off is important in medical diagnosis.

2. Consider the following Hidden Markov Model:

States: S = {H, C} (H = Hot, C = Cold)

Observations: O = {1, 2}

Initial probabilities: p_0(H)=0.5, p_0(C)=0.5

Transition probabilities:

P(H→H)=0.6, P(H→C)=0.4

P(C→H)=0.5, P(C→C)=0.5

Emission probabilities:

P(1|H)=0.8, P(2|H)=0.2

P(1|C)=0.3, P(2|C)=0.7

Observed sequence:

O=(1,2)

(1) (4%) Fill in the Viterbi table.

State\Time	***t*=0**	***t*=1 (O=1)**	***t*=2 (O=2)**
Hot (H)	0.5		
Cold (C)	0.5		

(2) (3%) Determine the most likely hidden state sequence.

(3) (3%) Why is Viterbi more efficient than checking all possible state sequences?

3. A processor has the following specifications:
   - Base CPI = 1.2
   - Clock rate = 2.5 GHz
   - Instruction mix: 40% ALU, 30% Load, 20% Store, 10% Branch
   - Cache specifications:
     - L1 I-Cache: 95% hit rate, 2-cycle miss penalty
     - L1 D-Cache: 90% hit rate, 8-cycle miss penalty
     - L2 Cache (unified): 98% hit rate, 50-cycle miss penalty

   (1) (6%) Calculate the effective CPI considering cache misses for load/store instructions. Assume all cache misses go to L2, and L2 misses go to main memory.

   (2) (4%) If a program executes 2 billion instructions, what is the total execution time?

4. A company deploys a web server behind a stateful firewall and NAT (Network Address Translation) device. The internal web server IP address is 192.168.1.10. The NAT device has a public IP address 140.113.10.5. The firewall is configured with DNAT (Destination NAT), forwarding traffic aimed at 140.113.10.5:8080 to the internal server at Port 80.

   (1) (2%) Explain the role of the NAT table.

   (2) (4%) An external client (IP: 8.8.8.8) on the Internet accesses the website using a browser. The client's OS assigns a random source port of 55000. Please fill in the packet header information after it has been processed by the NAT device as it enters the internal network, and as the response from the server passes through the firewall back to the Internet. Fill in the following table (e.g. 127.0.0.1:80).

Direction (after NAT)	Source IP:Port	Destination IP:Port
Client→server		
Server→client		

   (3) Assume the web server uses HTTP. Explain how a Man-in-the-Middle attacker located on the same LAN as the server could intercept and modify HTTP traffic. Your answer should include:

   a. (2%) One specific attack technique.

   b. (2%) Why NAT does not prevent this attack?