國立中正大學
108學年度碩士班招生考試
試 題

[第2節]

系所組別	資訊工程學系－乙組
科目名稱	計算機概論(含程式設計)

－作答注意事項－

※作答前請先核對「試題」、「試卷」與「准考證」之系所組別、科目名稱是否相符。

1. 預備鈴響時即可入場，但至考試開始鈴響前，不得翻閱試題，並不得書寫、畫記、作答。
2. 考試開始鈴響時，即可開始作答；考試結束鈴響畢，應即停止作答。
3. 入場後於考試開始40分鐘內不得離場。
4. 全部答題均須在試卷（答案卷）作答區內完成。
5. 試卷作答限用藍色或黑色筆（含鉛筆）書寫。
6. 試題須隨試卷繳還。

國立中正大學 108 學年度碩士班招生考試試題

科目名稱：計算機概論(含程式設計)　　本科目共 2 頁　第 1 頁

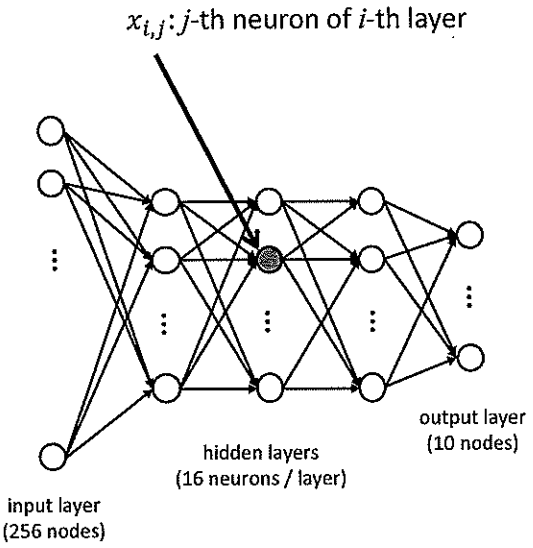
系所組別：資訊工程學系-乙組

1. (20pts) Quicksort is a basic sorting algorithm. (a) Please describe how it works. (b) Please analyze its time complexity in the worst case and in the average case.

2. (20pts) TCP has multiple congestion-control strategies. The followings are two basic mechanisms. Please describe their concepts: (a) congestion window (b) slow start.

3. (10pts) Please design a recursive version of a function to calculate the factorial of a positive integer n, i.e. $n!=1\times2\times3\times...\times n$.

4. (20pts) The following figure depicts a fully-connected neural network for recognizing handwritten digits (i.e. 0, 1, 2, … , 9), which contains 1 input layer, 3 hidden layers, and 1 output layer.



- The input layer has 256 nodes, each of which represents an 8-bit pixel of a 16*16 grayscale image.
- Each of the 3 hidden layers contains 16 neurons, which computes a weighted sum of all nodes in its precedent layer, adds a bias value, and performs ReLU activation (i.e. the result remains the same if it has a positive value; otherwise the result becomes 0). In other words, the j-th neuron of the i-th layer: $x_{i,j}$ computes $\max\left(\sum_{k=0}^{N-1} x_{i-1,k} \times w_{i,j,k} + b_{i,j} \quad 0\right)$, where N=256 for i=1 and N=16 for i=2 & 3.
- The output layer has 10 nodes, each of which represents a digit (i.e. 0, 1, 2, … , 9). The j-th output node $x_{4,j}$ computes $\sum_{k=0}^{15} x_{3,k} \times w_{4,j,k} + b_{4,j}$ without ReLU. The output with the maximum value will be the inference result.

(a) Write a C program to compute the 10 outputs (represented in IEEE single-precision floating-point format) from the 256 inputs (represented in unsigned char). Assume weights and biases are represented as single-precision floating-point numbers.

(b) How many weights and biases are needed respectively to compute the outputs of the neural network shown above for one 16*16 image?

5. (30%) Due to cost issues, a practical system usually implements only a fraction of required memory on chip and manages on-chip and off-chip data movements with the cache mechanism.
   (a) Assume there is only a 1KByte SRAM on chip for the weights in Problem 4 (assume biases are handled independently), and a design team decides to implement a direct-mapped cache to manage this 1Kbyte on-chip SRAM (all weights are stored off chip). Assume one cache block stores 32-byte data. What are the on-chip storage required in addition to the 1Kbyte SRAM to implement the direct-mapped weight cache? (Hint: cache tag …)
   (b) What is the miss rate of the weight cache in (a)? What is the type of cache miss (i.e. compulsory, conflict, or capacity)?
   (c) Describe an effective method to improve the weight memory organization in (a) under the same cost constraint.