系所組別：醫學資訊研究所

考試科目：程式設計 考試日期：0223，節次：2

※ 考生請注意：本試題不可使用計算機

Part I.

演算法 (50%)

一、 是非題(20 points) For each of the following statements, determine whether it is correct (T) or not (F).

1. The solution to the recurrence $T(n)=7T(n/2)+n^3$ is $T(n) = O(n^3 lgn)$.
2. Breadth-first search runs asymptotically faster than the depth-first search in a sparse graph.
3. The factional knapsack problem can have the optimal answer by the greedy strategy.
4. Sorting 8 elements with a comparison sort requires 24 comparisons in the worst case.
5. A problem is polynomial-time solvable if there exists an algorithm to solve it in time $O(n^k)$ for some constant k.

二、 計算題

1. (15 points) Give a **tight** asymptotic bound for $T(n)$, where $T(n) = T(2^{\sqrt{\lg n}}) + 1$.
2. (15 points) What is an optimal Huffman code for characters (A, G, C and T) in the DNA sequence **ACGTCCCAAGGTCAATCGCGGCGCGCGCG?** Please also compare the encoding costs between using the fixed length encoding and variable length encoding by your Huffman code.

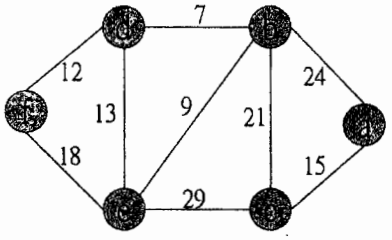
(背後仍有題目，請繼續作答)

※ 考生請注意：本試題不可使用計算機

Part II.

資料結構 (50%)

1. (15 points) Consider the following graph:



Compute minimum cost and construct minimum spanning tree step by step by using:

(1) Kruskal's algorithm,

(2) Prim's algorithm,

(3) Sollin's algorithm.

2. (15 points) Jeremy is a waiter working in a restaurant. The chef there is sloppy; when he prepares a stack of pancakes, they come out all different sizes. When Jeremy delivers the pancakes to the customer, he wants to rearrange them by grabbing several from the top and flipping them over on the way. After repeating this for several times, the smallest pancake is on top, and so on, down to the largest at the bottom. If there are n pancakes, how many flips are required? Design an algorithm to help Jeremy, and analyze its time complexity.

3. (20 points) A *Bloom Filter* is a space-efficient probabilistic data structure used to test whether a key is in a large data set. Instead of answering "yes" or "no," a Bloom Filter answers "maybe" or "no." A Bloom Filter consists of m bits of memory and h uniform and independent hash functions $f_1, f_2, \ldots, f_h$. Each f_i hashes a key k to an integer in the range $[1, m]$. Initially all m filter bits are zero, and the data set is empty. When key k is added to the data set, bits $f_1(k), f_2(k), \ldots, f_h(k)$ of the filter are set to 1. When a query "Is key k is in the data set?" is made, bits $f_1(k), f_2(k), \ldots, f_h(k)$ are examined. The query answer is "maybe" if all these bits are 1. Otherwise, the answer is "no."

(1) When the answer is "no," the key **is not** in the data set; when the answer is "maybe," the key **may or may not** be in the data set. Explain why.

(2) A *filter error* occurs whenever the answer is "maybe" and the key is not in the data set. Assume that key k is an integer in the range $[1, n]$ and u updates are made. Compute the probability of filter error for an arbitrary query after the u-th update.