Final

Exam Date and Time

Thursday, June 13, 1996. 2:20PM–4:50PM.

Note

This is a closed-book exam. However, you may consult the A4-sized sheet of notes that you prepared in advance. Each problem accounts for 10 points, unless otherwise marked.

Problems

1. To the question "What's the difference between philosophy and science?", Russell replied as follows:

   > Well, roughly, you'd say science is what we know and philosophy is what we don't know. That's a simple definition and for that reason questions are perpetually passing over from philosophy into science as knowledge advances.

   Your task is to compress the text of Russell's answer using Huffman's encoding (treating each "end of line" as a blank space). Show the Huffman code tree and calculate the rate of compression (compared to 8-bit ASCII encoding). (15 points)

2. Explain why the time complexity of the KMP algorithm is $O(n)$, where n is the length of string A. (5 points)

3. Given two strings $abbc$ and $babb$, compute the minimal cost matrix $C[0..4, 0..4]$ for changing the first string character by character to the second one. Show the detail of your calculation for $C[4, 4]$.

4. Trees are the only type of connected undirected graphs that contain a vertex v such that there exists a DFS tree rooted at v with the same set of edges as some BFS tree rooted at v. Why?

5. Prove that if the costs of all edges in a given connected graph are distinct, then the graph has an unique minimum-cost spanning tree.

6. Design an algorithm to color a graph with two colors (such that no two adjacent vertices have the same color) or determine that such coloring is not possible.

7. (a) Prove that Floyd's *All_Pairs_Shortest_Paths* algorithm works correctly for graphs with negative weights as long as there are no negative-weight cycles.
   (b) Modify Floyd's algorithm to detect the existence of negative-weight cycles. (Hint: think about the value of $Weight[i, i]$).

8. (a) What is the significance of proving either "P = NP" or "P $\neq$ NP" ?
   (b) To prove "P = NP", it suffices to show that some NP-complete problem is in P. Why?

9. The Hitting_Set problem is as follows.

   > Given a collection C of subsets of a set S and a positive integer k, does S contain a hitting set for C of size k or smaller, that is, a subset $S' \subseteq S$ with $|S'| \leq k$ such that S' contains at least one element from each subset in C?

   Prove that the Hitting_Set problem is NP-complete. (Hint: add an additional restriction that the size of each subset in C is two and reduce a known NP-complete problem to the restricted problem.)

10. The Subgraph_Isomorphism problem is as follows.

    > Given two graphs $G(V_1, E_1)$ and $H(V_2, E_2)$, does G contain a subgraph that is isomorphic to H? (Two graphs are isomorphic if there exists a one-one correspondence between the sets of vertices of the two graphs that preserve adjacency.)

    Prove that the Subgraph_Isomorphism problem is NP-complete? (Hint: reduce a known NP-complete problem to the restricted version of the problem where H is a complete graph.)