

CIS 1210 — Data Structures and Algorithms

Homework Assignment 5

Assigned: February 18, 2025

Due: February 24, 2025

Note: The homework is due **electronically on Gradescope** on February 24, 2025 by 11:59 pm ET. For late submissions, please refer to the Late Submission Policy on the [course webpage](#). You may submit this assignment up to 2 days late.

- A. Gradescope:** You must select the appropriate pages on Gradescope. Gradescope makes this easy for you: before you submit, it asks you to associate pages with the homework questions. Forgetting to do so will incur a 5% penalty, which cannot be argued against after the fact.
- B. L^AT_EX:** You must use the [LaTeX template](#) provided on the course website, or a 5% penalty will be incurred. Handwritten solutions or solutions not typeset in LaTeX will not be accepted.
- C. Solutions:** Please write concise and clear solutions; you will get only a partial credit for correct solutions that are either unnecessarily long or not clear. Please refer to the [Written Homework Guidelines](#) for all the requirements. Ed will also contain a complete sample solution in a pinned post.
- D. Algorithms:** Whenever you present an algorithm, your answer must include 3 separate sections. Please see Ed for an example complete solution.
 1. A precise description of your algorithm in English. No pseudocode, no code.
 2. Proof of correctness of your algorithm
 3. Analysis of the running time complexity of your algorithm
- E. Collaboration:** You are allowed to discuss **ideas** for solving homework problems in groups of up to 3 people but *you must write your solutions independently*. Also, you must write on your homework the names of the people with whom you discussed. For more on the collaboration policy, please see the [course webpage](#).
- F. Outside Resources:** Finally, you are not allowed to use *any* material outside of the class notes and the textbook. Any violation of this policy may seriously affect your grade in the class. If you're unsure if something violates our policy, please ask.

1. [10 pts] MCCX vs MDC

The MCCX Space Agency is under attack! Aliens from the rival planet, MDC, are trying to steal valuable data on our newly discovered planets. Determined to protect this knowledge, Cosmic Zimo, decides to encrypt the planet names using Huffman encoding. He uses a set S of n symbols with frequencies f_1, f_2, \dots, f_n to create these encodings. However, the aliens are persistent. If they want access to MCCX's discoveries, they must show their worth by proving key properties of Huffman encodings.

- (a) What is the longest encoding Cosmic Zimo could have created in terms of n ? Give an example set of frequencies that would produce this case. Note that your set of frequencies must be defined in such a way that it is generalizable for any value of n . This set of frequencies must be **valid**, meaning that the frequencies of all symbols sums to 1, though you do not need to prove this. Additionally, you do not need to prove that your proposed set of frequencies will produce the desired result.
- (b) Treat this question independently from your answer to part(a). For some Huffman tree, we recursively label each node with a cumulative frequency c_x . We define c_x for some leaf node x as the actual frequency of the symbol represented by that leaf node. We define c_x for each internal vertex x as the sum of its childrens' cumulative frequencies.

Let N represent the sum of all nodes' cumulative frequencies. Prove that the Average Bit Length (ABL) is equal to $N - 1$.

Hint: Draw a few examples to gain some intuition. Once you're done with your solution, think about how this observation can aid your understanding of the Huffman tree and related proofs.

2. [15 pts] Galactic Energy Alignment

In the depths of the Algorithmic Nebula, the notorious anomaly Heap Goyal still wreaks havoc. In response, the legendary Captain Leah has devised a plan to defeat Heap Goyal by combining the energies of her crew with mystical cosmic relics. Each cadet i , of n total cadets in the fleet, carries a unique energy level l_i . Each relic j , of n total relics, possesses a distinct stellar power p_j . According to the ancient codex of the *Radiant Ambassadors of Justice and Intellectual Vitality* (RAJIV), there exists a sacred cosmic constant x . The codex proclaims that when the energy level of a cadet and the stellar power of a relic combine to equal x , the resulting energy surge is enough to dispel the menace of Heap Goyal once and for all. Therefore, Captain Leah wants to pair cadets with relics such that $l_i + p_j = x$.

Help Captain Leah create an algorithm that returns d such (cadet, relic) pairs. You may assume that all of the n cadets and relics can be put into valid, disjoint pairs that sum to x .

- (a) Give an algorithm that does this in $O(dn)$.
- (b) Give an algorithm that does this in $O(n + d \log n)$.

Important Note: For both parts, you only need to provide an algorithm and running time justification, but we recommend thinking about how you would go about proving its correctness.