**A Combined Use of Linked List and Binary Tree**

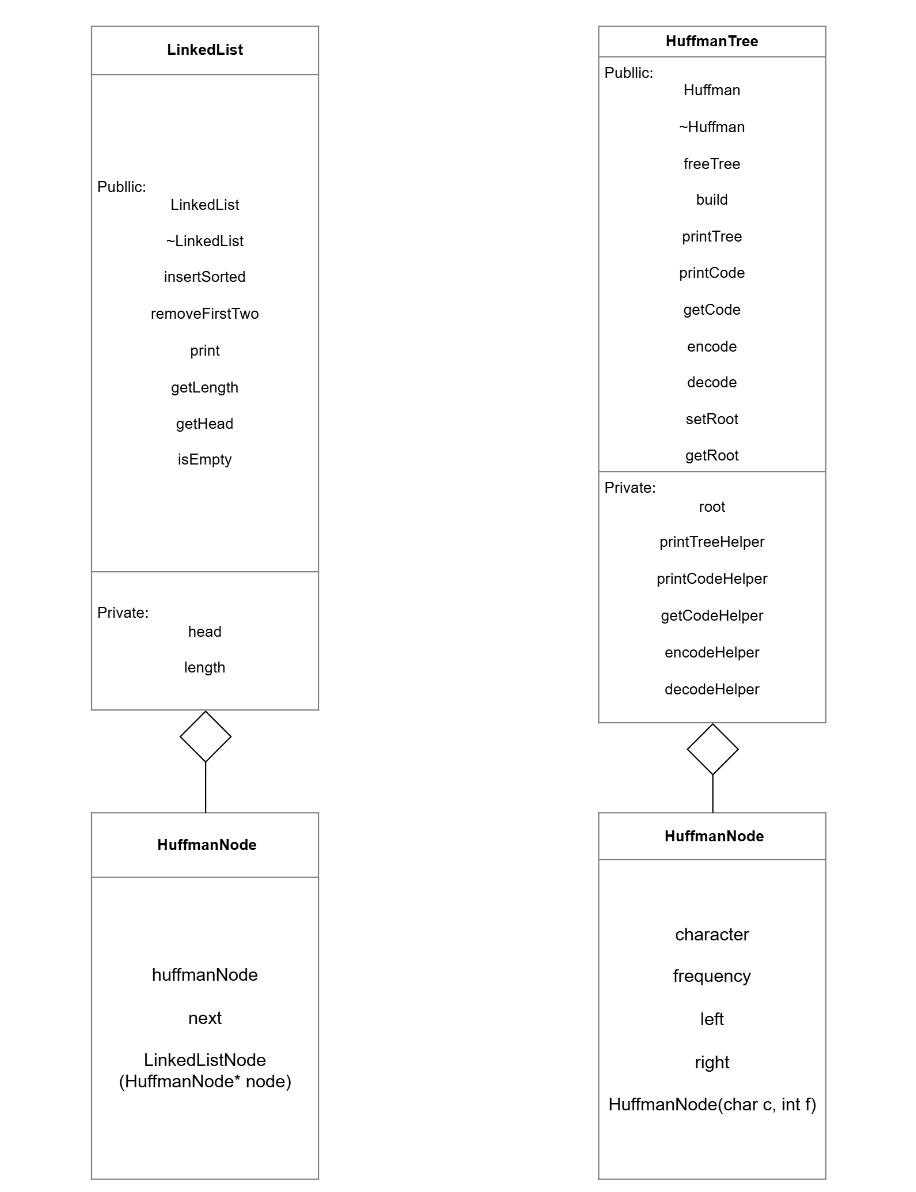
Group Number: Team Honor 3

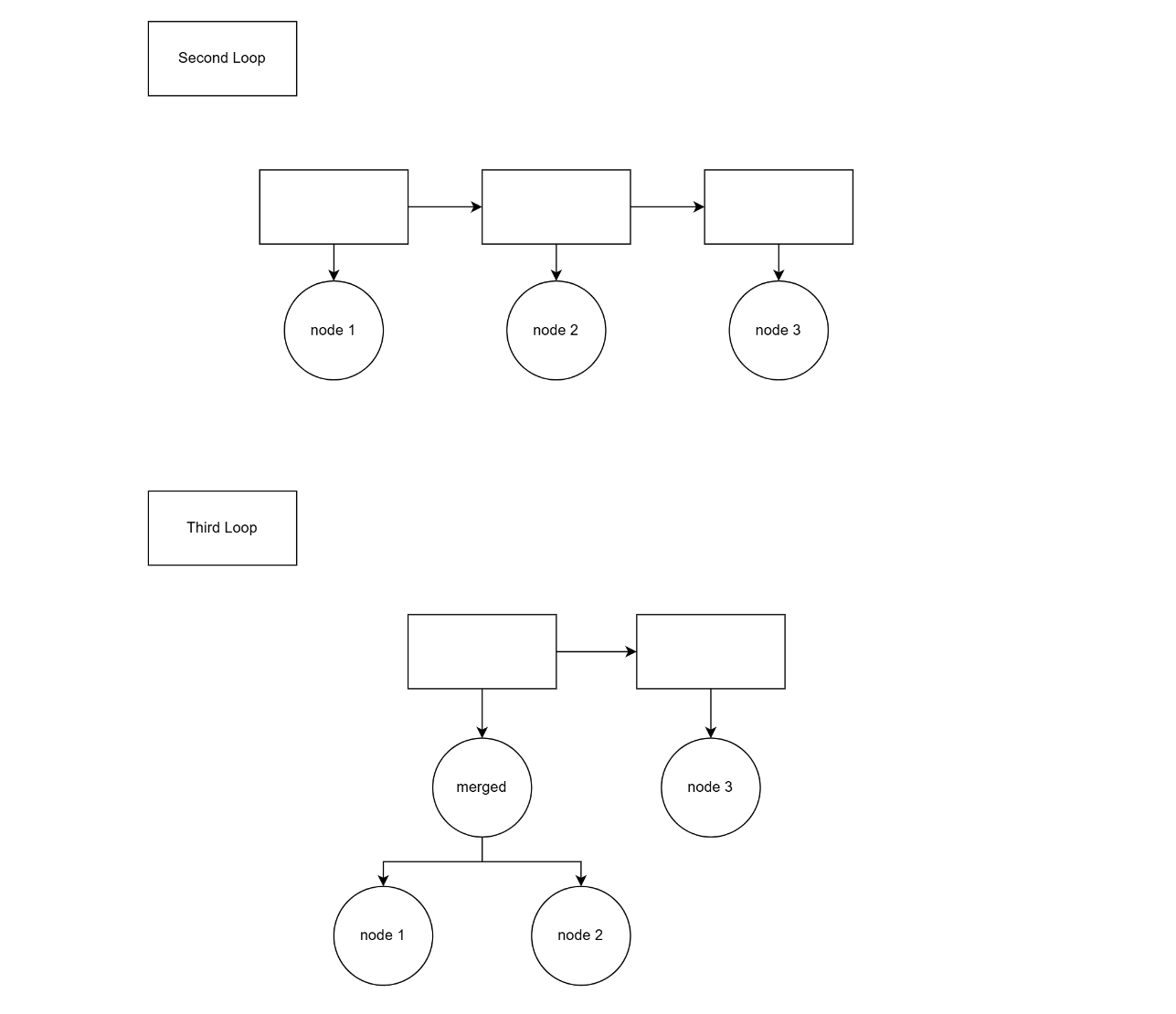
Team Members: Yankui Lee, Hengli Jia

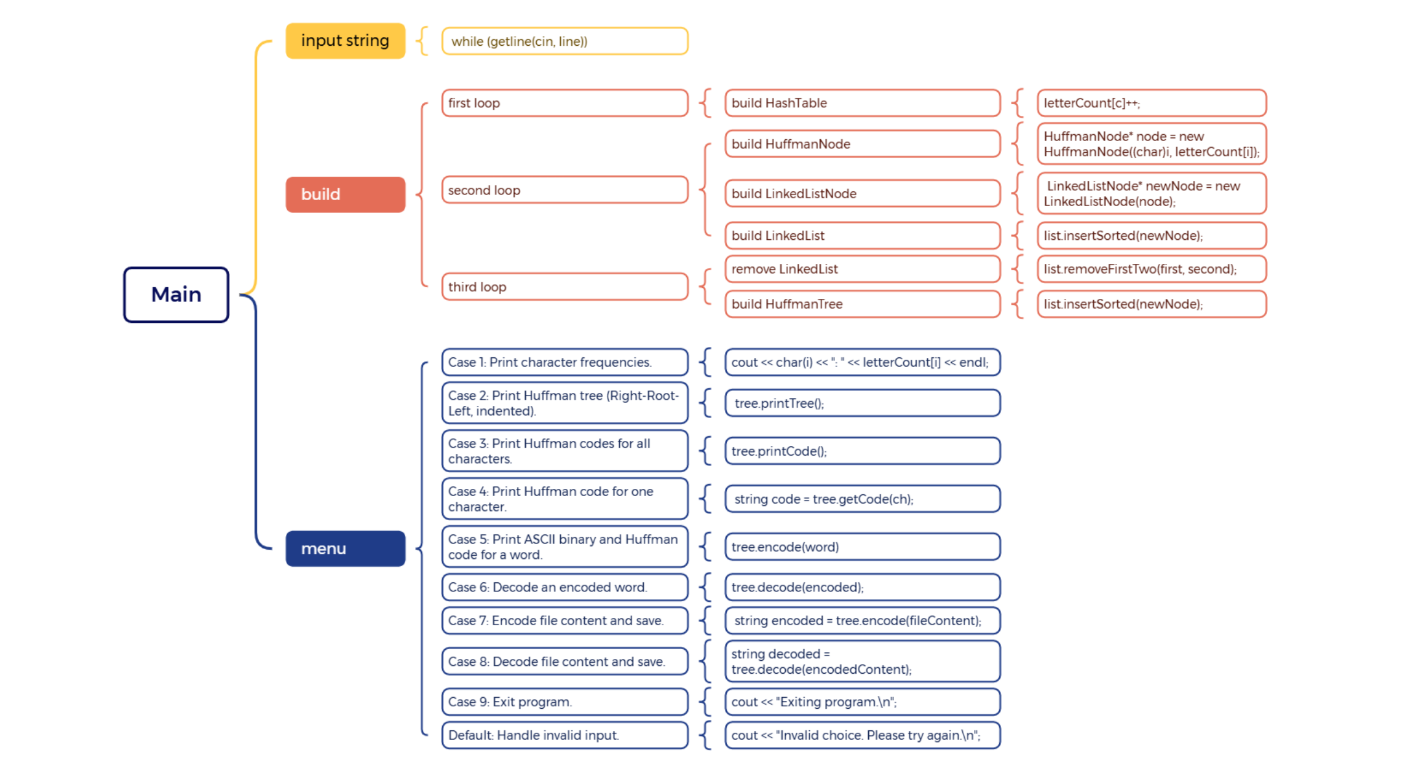
**Introduction**

Our project combines two data structures—linked lists and binary trees—to construct a Huffman tree. This tree structure is used to efficiently compress text by following a rule that minimizes storage space.

We first use the hash table concept to determine the frequency of each character. After we got each character and its frequency, we treated it as a node and built a binary search tree. The leaf nodes are the only nodes that contain characters. We use the two least frequent nodes to create a parent node that contains two nodes, but without a character inside. By continuing to do it, we can get a root node that only has the frequency, which is the length of the string. After completing the tree, we start from the root and go through the tree(0 for left, 1 for right) until we reach the leaf node and get the character. The higher the frequency, the easier the approach.

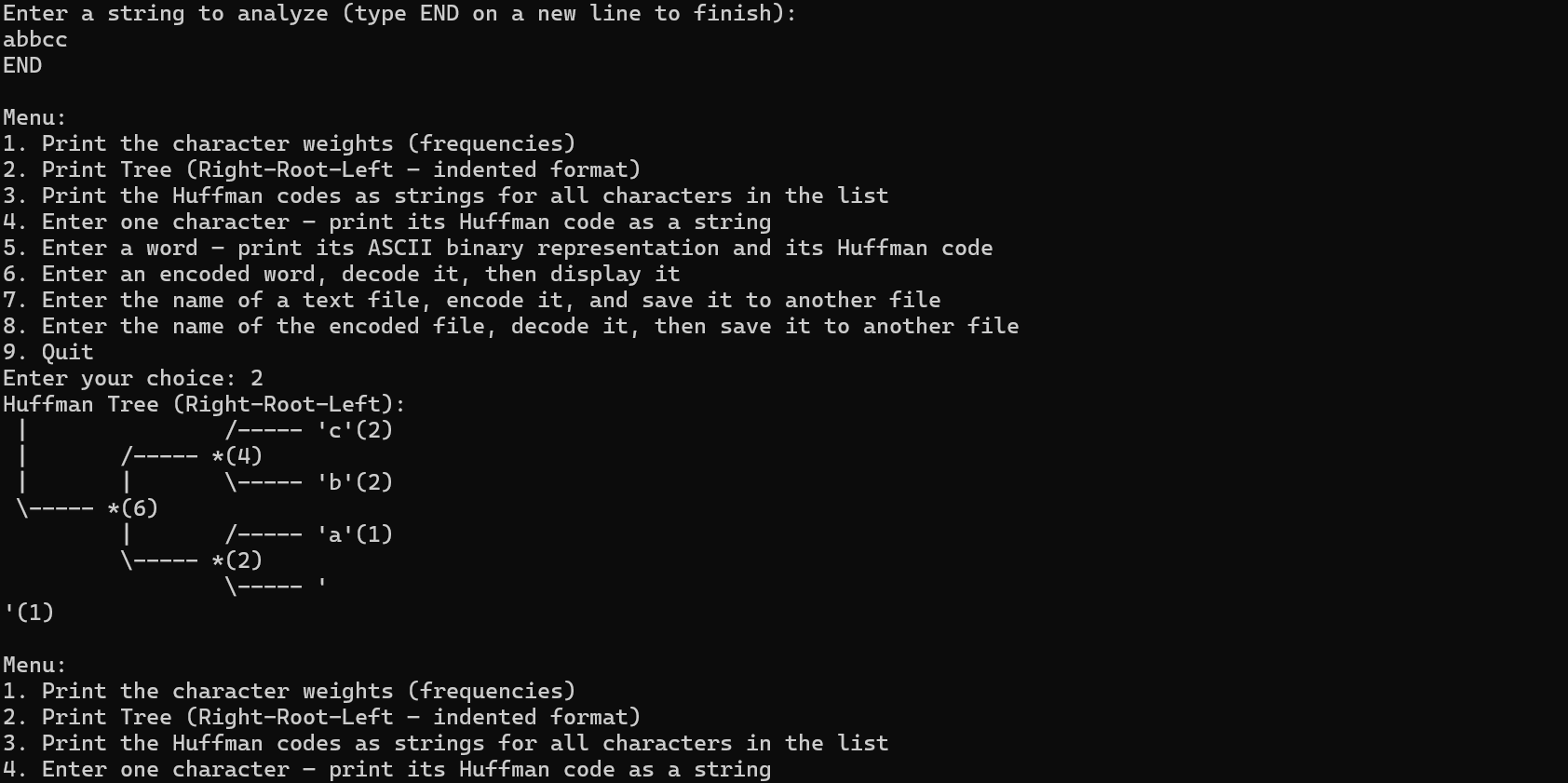
**Data Structures, Class Diagram, and Structure Chart**





**Sample Output Screen**

**Sample Screen Output**



**Major Debug Problems**

**First Problem**

The initial challenge we faced was not fully understanding how to structure the linked list. We rushed into implementation without a clear plan, resulting in a disorganized structure. This lack of preparation forced us to rebuild parts of the project, which consumed a significant amount of time.

Lesson Learned:  
To prevent similar issues in future projects, we should begin with a structure chart, a data structure diagram, and a class diagram. These visual tools help clarify the organization and implementation strategies for large projects, ensuring a smoother development process.

**Second Problem**

The second major issue involved a recursive dependency between Huffman.h and LinkedList.h. Specifically, Huffman.h required the LinkedList structure, while LinkedList.h also depended on elements from Huffman.h. This circular dependency resulted in compiler errors because C++ does not allow circular inclusions.

While forward declarations offer a temporary fix, the better long-term solution is to rethink the design. Instead of both headers relying on each other, we should redesign the structures to establish a clear containment relationship. For instance, one structure can contain the other, removing the need for mutual dependencies. This not only resolves the circular inclusion issue but also simplifies the design of the main function.

**Third Problem**

The third issue we encountered was the lack of a helper function to support our recursive method. Initially, we embedded all logic into a single recursive function, which made the parameters unclear and the function harder to understand and maintain. This also made it difficult to test and debug individual components of the recursion.

Lesson Learned:  
 In future projects, we should create a separate helper function for recursive algorithms. This helper function can handle the internal logic and parameter management, while the main function provides a cleaner interface. Doing so enhances code readability, separates concerns, and allows for easier debugging and extension.

**Application**

**Original text:**

**In computer science and information theory, a Huffman code is a particular type of**

**optimal prefix code that is commonly used for lossless data compression. The process of**

**finding and/or using such a code proceeds by means of Huffman coding, an algorithm**

**developed by David A. Huffman while he was a student at MIT, and published in the**

**1952 paper "A Method for the Construction of Minimum-Redundancy Codes"[1].**

**The output from Huffman's algorithm can be viewed as a variable-length code table for**

**encoding a source symbol (such as a character in a file). The algorithm derives this table**

**from the estimated probability or frequency of occurrence (weight) for each possible**

**value of the source symbol. As in other entropy encoding methods, more common**

**symbols are generally represented using fewer bits than less common symbols.**

**Huffman's method can be efficiently implemented, finding a code in time linear to the**

**number of input weights if these weights are sorted [2]. However, although optimal**

**among methods encoding symbols separately, Huffman coding is not always optimal**

**among all**

**compression methods - it is replaced with arithmetic coding [3] or asymmetric numeral**

**systems [4] if a better compression ratio is required**

**Encoded text:**

**100000001001011010111101011101111001000110011111100001100100101110110111100101011111111100111001010011110011000100101110100000111010111001101101010001011000111000111111010000010110011100011110011111000010010001101011010111110101110010110101111010100111111110011001001100111110111001011100000011011010111000111001001110000110001110110011100111111101010010110101001010111001001101101110101111001011011100100001111010110110100000110111010111101010011111111000111000101110011110011001001101011110101110111101101000101001010110011000011010011111001111001011101000001101001010100100010010010111101000100110100110111001101111101011110101110111100100001111010001000110101000100001000110000101101000111111101110010000101010111111101000100110101001011010100010110110001010011011000101011011100111001010011100000011010100000110000110100011000101011011100100000111011110001110011111010111101010011111111011100100001010101111111111110011010011010000110110011011101111101110010010011010100101111000010010001101011010111110101110010110101111010100110110001010110111100011110011100101100111100101011011010000001100011100011110101010010011111100010101111100101010111001111110011110100001101100110100000101101110001010011010011110111000100000100011000010010001101011010111110101110010110111000010001011010010111111010001111111011100000111010011001111100100001100011100111111001000111100111001111011100010110000000100010110111000111100111001010011110111001000111000011001001100100100011111100111100110001011000111000111110101000101011101000001010100000100101010111111011100101111110011111000011001010101011100010011011100010111110011100011010100111100101110100000110001110001111111010000000010100010010000110000000111011100110110101000101101010010111101110001010110001001101110100011111011000001111000001100111110011000110010100110111001010111101100110100000000101010011111101000101010100001011101010111001010100000100001010000010110100011111110101000011001111100100011001111001011000010101110111000010010001101011010111110101110010010101011010011001111001010110110100000011000111000111101110101110111001011010000111111100001010011011111110000111110011110011101001100111110000101001110000011001111000011001011111000001111001011110010101101001110001110101111010100111111110001101111000011001011111100101110100000010100111100101011110101001101100010101101110011111001001010000110000101111111110010010110011101100001101010010110010101100010000011101111000111001110100110011111010111100010111000001111011100111111000011001100010110011111001011011010010111101010110100010001100001011010001111111001111001010110110100000011000111000111101110100111111000001100001010111101001100011100010110010011000110111100001100101111010100010110000101011101110001110001111111011110100001101101110101110011111110011110111001000010101000010111100001011010010011000111011001101010000011001011000011111000000100001111110010101111011001101010010111101010101111011100011000000001111001010111111111001010110011100001111011010110110001001101010110111001011101000001101111011110111100011101110011010010001000110100001100101111010100000101001111001000011111111010100101111000111000111111100100101000011000010111111111001001011001110110000110101001000010001101110001000100110011000101101010001110001111100001101111001000110000101011100110110011011110010101111010100110110001010110111011101111100111000110101001101001110001111011101101000001111110101111010111011110110100010010100010010110011101100001101010010010011001110000111111010110111110010111100000111100101001010110011000001111111001000011110100111100100011111110011110000110100011000101011011100101111111110000111100001101000010110001101001100011100010111001011010010111101000100110101111010111011110110100010110010010110011101100001101010010010000010000101000001001000110101101011111010111001001010101101001101110111110011100011010100111101011101110010110100001111111011110101101011011010111011011110010001110010101100110011011101111001100101111111011111001000111111100111110001111001011011000101001101100010101101110011111010111101010011111111001100010110001101101110111111101001001100010111101110000110001110101100011100011111010100001000011111011000011111000011010100101111001100010111001000110011110111000011110110101101100010011010011001100101111000111000111110100111111011100001111011010110110001001101001100111000011111100100101000000011111110011110000101110101011110101010000010001100001001101011100001111000101011110000111000111100111100100011100011010000111011011000111010101110010011011011101011110010010100011111101101000101011011101110111110011100011010100110100110111100101011110101001101100010101101110010010110011101100001101010010010011001001111111001011100000111001111111001010110011100011110000100100011010110101111101011100101101011110101001101100010101101110011001001100010101000111100111100101110000011110110001001101010111001001101101110101111001001010001111110110100010101101110011110010100100101001011110101110111100100001111010001000110101000101101110111110011100011010100110100110100000111110011000111100110010011000001111111001100100111101111111100111101110000011000111000111001110000011000111000111101111100110110101111101011110101001101100010101101110000101111000000111010101001101010000011001110100101100111011110111110011000001101011111000100001111101111100000111100100101000100101100010000111111111010100110000101111000001000010101001100110010111100111110100001111100110011111100001101011110101110111100100001111010001000110101000101100000011100110110101011001100100110000011111000000100001101100000111110011**

**Assignment**

**Yankui Lee:**

Data Structure Design

main(), menu() and other related functions

Linked List algorithms

Integration, testing, and project documentation

**Hengli Jia:**

Data Structure Design

Binary Tree algorithms

Hash Table algorithms

Integration, testing, and project documentation

**Conclusions**

Our project successfully combined linked lists and binary trees to implement a Huffman tree for efficient text compression. By first using a hash table to calculate character frequencies, we created nodes representing these characters and organized them into a binary tree structure. The construction of the tree followed Huffman’s algorithm: repeatedly combining the two least frequent nodes into a new parent node until a single root node remained. Traversing the tree from root to leaf (using 0 for left and 1 for right) allowed us to assign binary codes to each character, with more frequent characters receiving shorter codes—thus optimizing compression.

Through this process, we faced and overcame two significant challenges. First, a lack of planning in our initial linked list structure led to a confusing and inefficient implementation. From this, we learned the importance of starting with visual planning tools such as structure charts, data structure diagrams, and class diagrams to guide development and avoid rework.

Second, we encountered a circular dependency between Huffman.h and LinkedList.h, causing compiler errors. This highlighted the importance of thoughtful design and modular code. While forward declarations can serve as a quick workaround, we learned that a cleaner solution is to establish a one-way containment relationship between structures to eliminate mutual dependencies and simplify the overall system architecture.

Ultimately, this project not only helped us understand how to implement Huffman coding using fundamental data structures, but also reinforced best practices in software design and project planning. These lessons will be valuable in future coding endeavors, especially when tackling complex, multi-module systems.