Recall the statement of the Minimum Cost Prefix Tree problem.

**Given:** A set $C$ of $n \geq 2$ symbols to be encoded; for each symbol $c \in C$, a frequency value $f(c)$.

**Required:** A codeword for each symbol in $C$ such that the encoding cost is minimized.

### I. Recursive Version of Huffman’s Algorithm:

1. **if** $(|C| = 2)$ **then**
   
   Encode one symbol using 0 and the other using 1.

   **else**
   
   (a) Let $x$ and $y$ be the symbols with the lowest frequencies in $C$.
   
   (b) Combine $x$ and $y$ into a new symbol $w$ with frequency $f(w) = f(x) + f(y)$.
   
   (c) Recursively construct a tree $T'$ for $C' = (C - \{x, y\}) \cup \{w\}$.
   
   (d) From $T'$, construct tree $T$ for $C$ by making $x$ and $y$ the children of $w$.

2. Return $T$.

### II. Iterative Version of Huffman’s Algorithm:

1. Create a node (of the tree) corresponding to each symbol in $C$. Let $S$ denote the resulting set of nodes.

2. **while** $(n \geq 2)$ **do**
   
   (a) Let $x$ be a node with smallest frequency value in $S$. Remove $x$ from $S$.
   
   (b) Let $y$ be a node with smallest frequency value in $S$. Remove $y$ from $S$.
   
   (c) Create a new node $w$ with $x$ and $y$ as its children. Add $w$ to $S$.
   
   (d) Set $f(w) = f(x) + f(y)$.
   
   (e) Set $n = n - 1$.

3. For each internal node of the tree, label the edge to the left child as 0 and the edge to the right child as 1.

4. For each leaf $c$ (i.e., symbol in $C$), construct its codeword by concatenating the bits in the path from the root to $c$. 

(over)
Note: The correctness of Huffman’s Algorithm (Theorem 4 below) can be shown through the proofs of the following three lemmas.

Lemma 1: (a) Every canonical binary tree represents a prefix code. (b) Each prefix code can be represented by a canonical binary tree.

Lemma 2: Every optimal solution to the problem is a full binary tree (i.e., each internal node has exactly two children).

Lemma 3: Let $x$ and $y$ be two symbols in $C$ with the two lowest frequencies. There is an optimal solution for $C$ in which $x$ and $y$ are leaves with maximum depth.

Theorem 4: Huffman’s Algorithm produces an optimal solution.