Отчет к заданию 10, Вариант 2. Никита Бочков 2254

**Знакомство с методом и кодами Хаффмана. Применение контейнеров STL. Доработка проекта для декодирования потока. Перестроение таблицы кодов – компактификация таблицы.**

**Как строится частотная таблица символов?**

Для построения частотной таблицы символов программа сначала проходит по всему входному файлу и подсчитывает количество вхождений каждого символа. Эта информация сохраняется в виде пары (символ, частота) в специальную таблицу. Таким образом, для каждого уникального символа определяется, сколько раз он встречается в тексте.

**Как строится дерево по частотам символов?**

Дерево Хаффмана строится следующим образом:

1. Каждый символ из частотной таблицы представляется как листовой узел дерева с весом, равным частоте этого символа.
2. Эти узлы добавляются в приоритетную очередь, отсортированную по частотам (весам) узлов.
3. Из очереди последовательно извлекаются два узла с наименьшими частотами. Эти узлы объединяются в новый узел, частота которого равна сумме частот двух объединенных узлов.
4. Новый узел добавляется обратно в очередь.
5. Процесс повторяется до тех пор, пока в очереди не останется только один узел, который становится корнем дерева Хаффмана. Это дерево используется для кодирования символов.

**Приведите исходные требования к кодам Хаффмана. Однозначно ли они определяют коды для заданной таблицы частот символов?**

Основные требования к кодам Хаффмана заключаются в следующем:

1. Коды должны быть префиксными, то есть ни один код не может быть префиксом другого кода. Это обеспечивает однозначную декодировку последовательности битов.
2. Символы с более высокой частотой вхождений получают более короткие коды, а символы с низкой частотой — более длинные. Это позволяет минимизировать среднюю длину кодированной последовательности.

Эти требования однозначно определяют коды для заданной таблицы частот символов. Однако, если несколько символов имеют одинаковую частоту, возможно несколько различных, но эквивалентных деревьев Хаффмана, что приводит к разным, но корректным наборам кодов.

Компрессия методом Хаффмана эффективно снижает размер данных при наличии повторяющихся символов, и обеспечивает уникальность декодирования благодаря префиксным кодам.