

Формат закодированного файла

Общий принцип

Закодированный файл представляет собой бинарный файл, содержащий закодированный текст. Файл начинается с метаданных, включающих информацию о структуре дерева кодирования и расширении файла.

Метаданные

Инициалы студента: StudentAlex.

Имя Оригинального Файла: [Оригинальное имя файла] Пример

test2.txt & LargeTest.txt

затем \n

Размер Таблицы: [Количество уникальных символов и их кодов в таблице]\n

Таблица Кодов:

[Символ1][Код1]\n

[Символ2][Код2]\n

...

[СимволN][КодN]\n

Далее идет закодированная информация но первый байт отражает количество дописанных в конец 0

Дополнительные Нули:

[Дополнительные нули для выравнивания]

Закодированные Данные:

[Бинарные данные закодированного файла]

Текстовая часть

После метаданных следует закодированный текст. Он содержит все символы исходного текста, но записанные с помощью уникального кода для каждого символа. Соответствие кода символу выполняется на основании дерева Хаффмана.