ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к программе для сжатия данных по методу Хаффмана ZIPMeHuffman

Программа работает в двух режимах:

1. Архивирование файла.

На вход подается бинарный файл. Данный файл разбивается на байты, после чего строится модель: сначала генерируется лист из 256 элементов, где i-ый элемент отображает количество раз, которое байт, являющийся двоичным представлением i, встречается в файле. Далее для удобства модель представляется в форме словаря, где ключи – частоты, а значения – листы, содержащие в качестве элементов все такие i, что байт, являющийся двоичным представлением i, встречается в файле с данной частотой.

Далее строится код Хаффмена. В сжатый файл с именем «\_имя\_входящего\_файла.zmh» код Хаффмена записывается следующим образом: записывается кодируемый символ, далее длина соответствующего кодового слова и само кодовое слово, после чего программа переходит к записи следующего кодируемого символа. После того, как кодируемые символы исчерпаны, программа записывает ‘\n’ и переходит к кодированию и записи данных архивируемого файла. Для последнего неполного байта добавляется необходимое до полноты количество нулей, после чего записывается ещё один байт, содержащий информацию о том, какое число нулей было добавлено. После этого запись прекращается.

Впоследствии полученный файл будет использоваться в режиме 2.

2. Разархивирование файла.

На вход подаётся файл с расширением zmh. Программа проверяет правильность формата данного файла и, если файл имеет некорректное расширение, то сообщает об ошибке и предлагает заново ввести имя файла. В случае успеха программа выполняет чтение кода, записанного в файл, а затем чтение закодированных данных заархивированного файла и их декодирование с помощью полученного кода. Полученный в результате разархивированный файл сохраняется под названием «decodedfrommh\_название\_исходного\_файла».