

Лабораторная работа № 2

Метод "Стопка книг"

Метод сжатия данных "Стопка книг" был впервые предложен Б. Я. Рябко и затем переоткрыт на Западе под названием Move-To-Front (MTF).

Идея метода состоит в следующем. Буква сообщения кодируется числом, представляющим ее текущий номер в алфавите источника. Затем она ставится в начало алфавита, при этом остальные буквы сдвигаются на одну позицию вправо. То есть буквы алфавита переупорядочиваются, с тем чтобы наиболее часто встречающиеся буквы находились ближе к началу.

Рассмотрим пример. Пусть алфавит источника $A = \{a, b, c, d, e\}$ и генерируется сообщение *baadaade...* Покажем, как меняется порядок букв и какие числа формируются по мере поступления символов от источника.

Сообщение		<i>b</i>	<i>a</i>	<i>a</i>	<i>d</i>	<i>a</i>	<i>a</i>	<i>d</i>	<i>e</i>
Позиции букв в алфавите	0	<i>a</i>	<i>b</i>	<i>a</i>	<i>a</i>	<i>d</i>	<i>a</i>	<i>a</i>	<i>d</i>
	1	<i>b</i>	<i>a</i>	<i>b</i>	<i>b</i>	<i>a</i>	<i>d</i>	<i>d</i>	<i>a</i>
	2	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>b</i>	<i>b</i>	<i>b</i>	<i>b</i>
	3	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>
	4	<i>e</i>	<i>e</i>	<i>e</i>	<i>e</i>	<i>e</i>	<i>e</i>	<i>e</i>	<i>e</i>
Код		1	1	0	3	1	0	1	4

Последовательность 11031014 еще не является "кодом": каждое из входящих в нее чисел необходимо закодировать.

Сжатие достигается за счет того, что чаще встречающиеся буквы имеют меньшие номера позиций, которые кодируются более короткими кодами.

Способ декодирования очевиден.

Задание на ЛР. Разработать программу, которая с помощью метода "Стопка книг" и кодирования целых чисел сжимает заданный файл. В качестве буквы алфавита выступает произвольный байт.

Аудиторное время на выполнение и защиту работы – 4 часа.