# Учреждение образования

«Белорусский государственный технологический университет»

Факультет Информационных Технологий

Кафедра Информационных Систем и Технологий

Отчет по лабораторной работе № 10

«Сжатие/распаковка данных методом Лемпеля − Зива»

Выполнил: Шедько Евгений

студент 3 курса 3 группы

факультета ИТ

Проверил: Берников В.О.

# Цель:

# Приобретение практических навыков использования метод Лемпеля − Зива (Lempel-Ziv) для сжатия/распаковки данных.

# Задачи:

# Закрепить теоретические знания по алгебраическому описанию и использованию методов сжатия/распаковки (архивации/ разархивации) данных на основе метода Лемпеля − Зива.

# Разработать приложение для реализации метода Лемпеля − Зива.

# Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

# Теоретическая часть

В 1977 г. Авраам Лемпель и Якоб Зив выдвинули идею формирования «словаря» общих последовательностей анализируемых (сжимаемых) данных. При этом сжатие данных осуществляется за счет замены записей соответствующими кодами из словаря. Классический алгоритм Лемпеля − Зива – LZ77, названный так по году представления метода, формулируется следующим образом: «если в проанализированном (сжатом) ранее выходном потоке уже встречалась подобная последовательность байт, причем запись о ее длине и смещении от текущей позиции короче, чем сама эта последовательность, то в выходной файл записывается ссылка (смещение, длина), а не сама последовательность».

Суть метода LZ77 (как и последующих его модификаций) состоит в следующем: упаковщик постоянно хранит некоторое количество последних обработанных символов в буфере. По мере обработки входного потока вновь поступившие символы попадают в конец буфера, сдвигая предшествующие символы и вытесняя самые старые. Размеры этого буфера, называемого также скользящим словарем (англ. sliding dictionary), варьируются в разных реализациях систем сжатия.

Скользящее окно имеет длину n, т. е. в него помещается n символов, и состоит из двух частей:

1. последовательности длины n1 = n − n2 уже закодированных символов (словарь);
2. упреждающего буфера (буфера предварительного просмотра, lookahead) длиной n2 – буфера кодирования.

На рис. 1 изображено кодируемое сообщение и метод кодирования Лемпеля-Зива, представляющий из себя выделения памяти для словаря и буфера, в которые будут записываться символы, а также поиск первых символов из буфера в словаре и запись результата поиска как код, состоящий из номера бита повторяющейся подстроки (p), длины этой повторяющейся строки (q) и буквы, следующей за повторяющейся строкой в буфере, после чего будет произведён сдвиг на длину повторяющейся строки + 1 (q + 1):

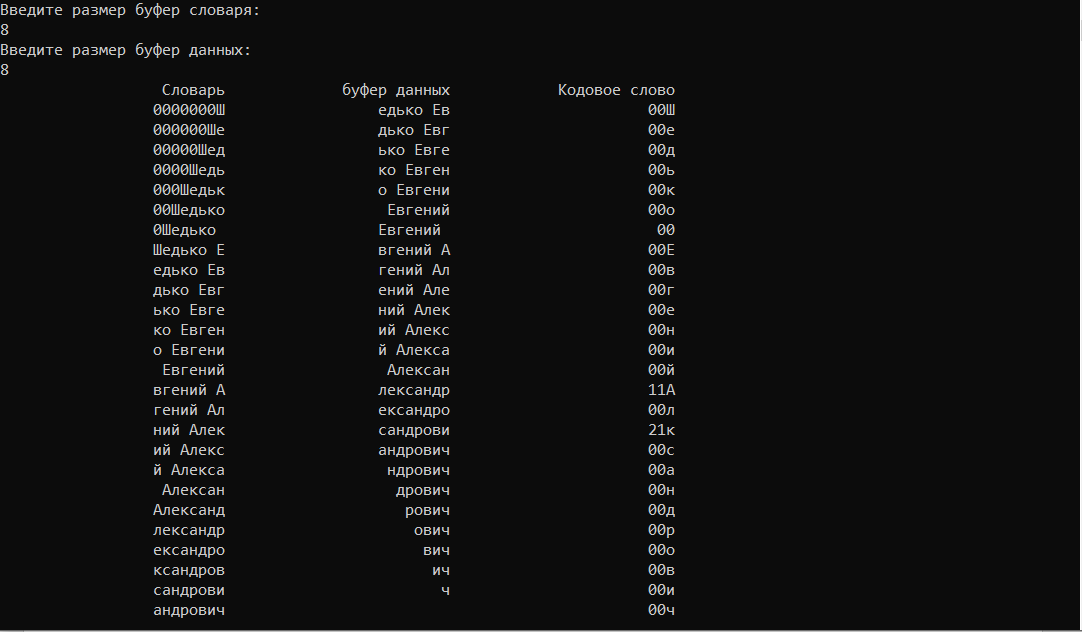
****

Рис.1 – Кодируемое сообщение и кодирование Лемпеля-Зива

На рис. 2 изображено сообщение после кодирования методом Лемпеля-Зива, состоящее из последовательной записи кодов каждого шага алгоритма на рис. 1:

****

Рис.2 – Сообщение после кодирования

На рис. 3 изображён алгоритм декодирования, обратный алгоритму, представленному на рис. 1:

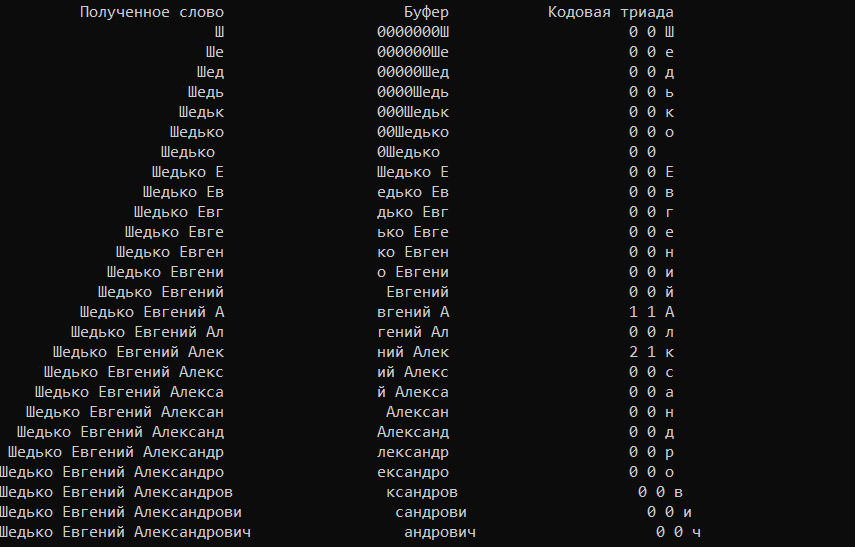
****

Рис.3 – Декодирование сообщения, закодированного методом Лемпеля-Зива

# Вывод: в данной лабораторной работе мною были приобретены навыки использования метода Лемпеля − Зива (Lempel-Ziv) для сжатия/распаковки данных.