БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

КАФЕДРА ИНФОРМАЦИОННЫХ СИСТЕМ И ТЕХНОЛОГИЙ

**Лабораторная работа №8**

**«Сжатие/распаковка данных**

**методом Барроуза–Уилера**»

**Выполнила:**

студентка 3 курса 1 группы

Потапейко Полина Павловна

**Проверил:**

Берников Владислав Олегович

Минск 2021

**Цель:** приобретение практических навыков использования метода Барроуза–Уилера для сжатия/распаковки данных.

**Ход работы**

**Задание 1а.** Выполнить прямое и обратное преобразование блока данных, состоящего из собственного имени.

Первым шагом прямого преобразования является следующее: сообщение записывается в матрицу размерностью kxk, в каждой последующей строке происходит сдвиг сообщения на один символ влево (рис. 1).

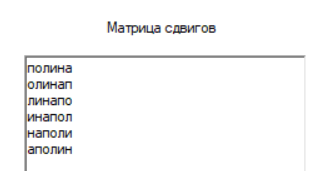


Рисунок 1 – Матрица сдвигов

Затем производится лексикографическая сортировка матрицы (сортировка по алфавиту) (рис. 2).

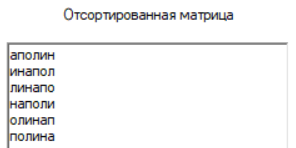


Рисунок 2 – Сортировка матрицы сдвигов

На выходе будет сообщение, состоящее из символов последнего столбца отсортированной матрицы, и число, являющееся номером строки исходного сообщения в отсортированной матрице (рис. 3).

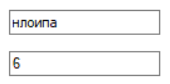


Рисунок 3 – Выходное сообщение

В случае обратного преобразования на вход подается сообщение и число, сообщение записывается в крайний правый свободный столбец матрицы той же размерности, что и при прямом преобразовании, затем производится лексикографическая сортировка строк матрицы, снова записывается сообщение в крайний правый свободный столбец, производится сортировка, и так k шагов, по числу i получают номер строки с исходным сообщением, если не произошло ошибок при передаче (рис. 4).

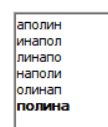


Рисунок 4 – Матрица обратного преобразования

**Задание 1б.** Выполнить прямое и обратное преобразование блока данных, состоящего из собственной фамилии.

Аналогично первому заданию выполняем те же действия с сообщением, состоящим из фамилии.

Получаем матрицу сдвигов (рис. 5).

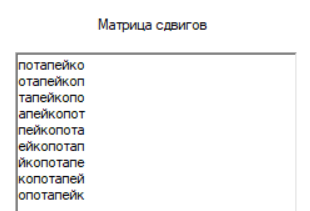


Рисунок 5 – Матрица сдвигов

Производим лексикографическую сортировку (рис. 6).

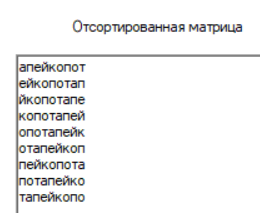


Рисунок 6 – Сортировка

Получаем выходное сообщение (рис. 7).

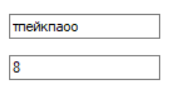


Рисунок 7 – Выходное сообщение

Получаем матрицу обратного преобразования (рис. 8).

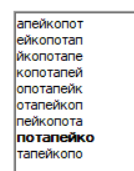


Рисунок 8 – Матрица обратного преобразования

**Задание 1в.** Выполнить прямое и обратное преобразование блока данных, состоящего из слова по варианту (вариант 10).

Аналогично выполняем те же действия с сообщением, состоящим из слова по варианту.

Получаем матрицу сдвигов (рис. 9).

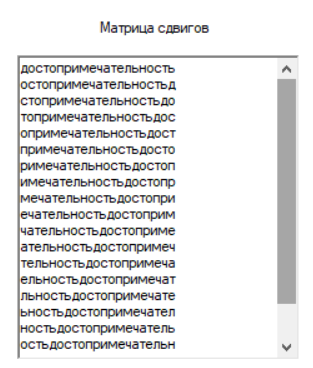




Рисунок 9 – Матрица сдвигов

Производим лексикографическую сортировку (рис. 10).

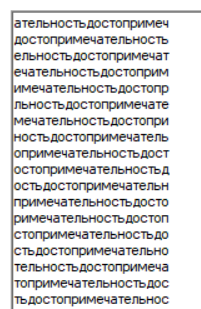




Рисунок 10 – Сортировка

Получаем выходное сообщение (рис. 11).

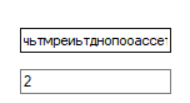


Рисунок 11 – Выходное сообщение

Получаем матрицу обратного преобразования (рис. 12).



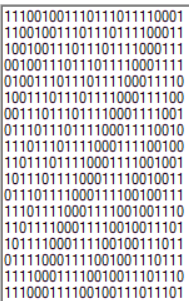
Рисунок 12 – Матрица обратного преобразования

**Задание 2.** Перевести первые 3 символа из блока данных своего варианта (вариант 10) в бинарную последовательность в соответствии с кодами ASCII. Выполнить прямое и обратное преобразование.

дос = 111001001110111011110001

Аналогично выполняем те же действия с сообщением, состоящим из бинарного кода.

Получаем матрицу сдвигов (рис. 13).



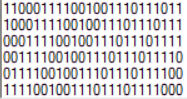
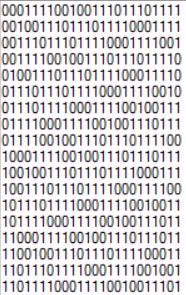


Рисунок 13 – Матрица сдвигов

Производим лексикографическую сортировку (рис. 14).



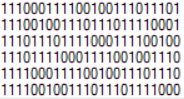


Рисунок 14 – Сортировка

Получаем выходное сообщение (рис. 15).

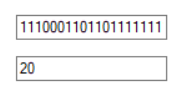
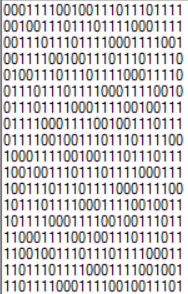


Рисунок 15 – Выходное сообщение

Получаем матрицу обратного преобразования (рис. 16).



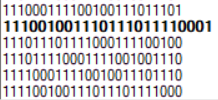


Рисунок 16 – Матрица обратного преобразования

**Вывод:** BWT-преобразование преобразует блок данных в формат, подходящий для дальнейшего сжатия. Прямое преобразование выполняется в 4 этапа: 1) выделение блока данных, 2) составление таблицы размером kxk всех циклических сдвигов входной строки, 3) лексикографическая сортировка строк таблицы, в результате чего получается таблица того же размера, 4) в качестве выходной строки выбирается последний столбец второй таблицы преобразования и номер строки, совпадающей с исходной строкой. Выходная строка всегда по объему превышает входную. Обратное преобразование заключается в выполнении k одинаковых шагов, каждый из которых состоит из двух операций, с целью воссоздания второй матрицы прямого преобразования: 1) в крайний справа пустой столбец матрицы записывается последовательность символов на входе, 2) лексикографическая сортировка столбцов заполненной части воссоздаваемой матрицы. После k шагов матрица будет получена. Число на входе будет указывать на позицию исходного сообщения в воссозданной матрице.