УO «Белорусский государственный технологический университет»

Кафедра информационных систем и технологий

Лабораторная работа

**«Метод сжатия Лемпеля-Зива»**

Выполнил студент

3 курса, 4 группы

Чаевский А. В.

Проверила

Герман Ю. О.

Минск 2019

# **1 Постановка задачи**

Цель: Самостоятельно освоить технику компрессии файлов по методу Лемпеля-Зива.

# **2 Краткие сведения**

Рассмотрим алгоритм Лемпеля-Зива на примере.

Пусть дана последовательность: aabcdabcddaabcddacd

Строим таблицу:

|  |  |  |
| --- | --- | --- |
| символ | Следующий символ | Кодировка |
|  |  |  |

Первый символ – a. Его в таблице (в левом столбце) нет. Добавляем его в таблицу:

abcdabcddaabcddacd

|  |  |  |
| --- | --- | --- |
| Символ | Следующий символ | Кодировка |
| a | a |  |
|  |  |  |

Следующий символ – а. В таблице он уже есть. За ним следует символ b. Получаем новую комбинацию символов – ab, которой в таблице еще нет. Помещаем ее в таблицу:

cdabcddaabcddacd

|  |  |  |
| --- | --- | --- |
| Символ | Следующий символ | Кодировка |
| a | a |  |
| ab | c |  |

Продолжаем с символа c. Такого символа в таблице нет. Помещаем его в таблицу и выписываем следующий символ:

dabcddaabcddacd

|  |  |  |
| --- | --- | --- |
| Символ | Следующий символ | Кодировка |
| a | a |  |
| ab | c |  |
| c | d |  |

Следующий символ – d. В таблице его нет. Помещаем в таблицу:

abcddaabcddacd

|  |  |  |
| --- | --- | --- |
| Символ | Следующий символ | Кодировка |
| a | a |  |
| ab | c |  |
| c | d |  |
| d | a |  |

Cледующий символ – a. В таблице он есть. За ним идет символ b. Получаем комбинацию ab. Она также есть в таблице. За ab идет символ с. Получаем комбинацию abc. Такой комбинации нет. Включаем ее в таблицу и получаем

abcddaabcddacd

|  |  |  |
| --- | --- | --- |
| Символ | Следующий символ | Кодировка |
| a | a |  |
| ab | c |  |
| c | d |  |
| d | a |  |
| abc | d |  |

Продолжаем далее по аналогии и получаем следующую результирующую таблицу

ddaabcddacd

|  |  |  |
| --- | --- | --- |
| символ | Следующий символ | Кодировка |
| a | a | 0001 |
| ab | c | 0010 |
| c | d | 0011 |
| d | a | 0100 |
| abc | d | 0101 |
| dd | a | 0110 |
| aa | b | 0111 |
| b | c | 1000 |
| cd | d | 1001 |
| da | c | 1010 |

В результате входная последовательность aabcdabcddaabcddacd преобразуется к такому виду:

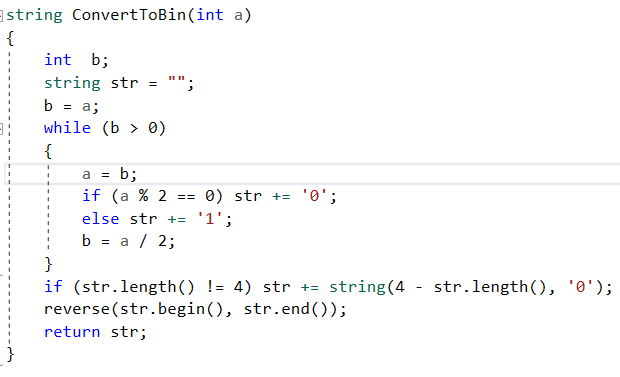
0001 0101 0100 0101 0110 0111 1000 1001 1010 1001

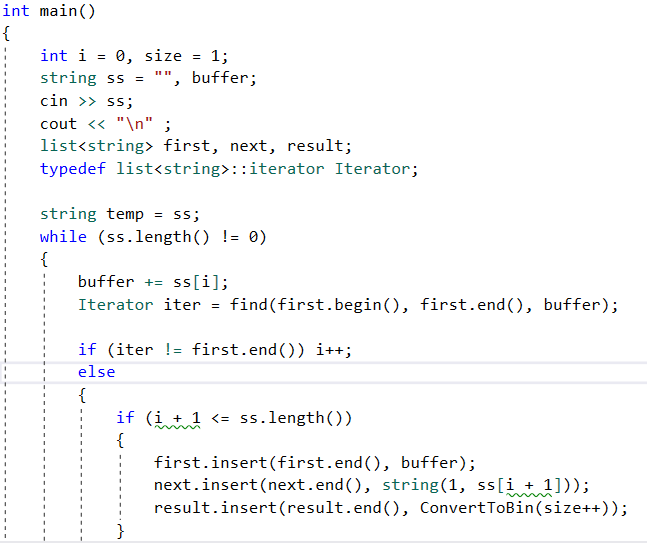
Это – двоичная последовательность. Она содержит 40 битов. А какова же последовательность битов исходной комбинации? Она такая

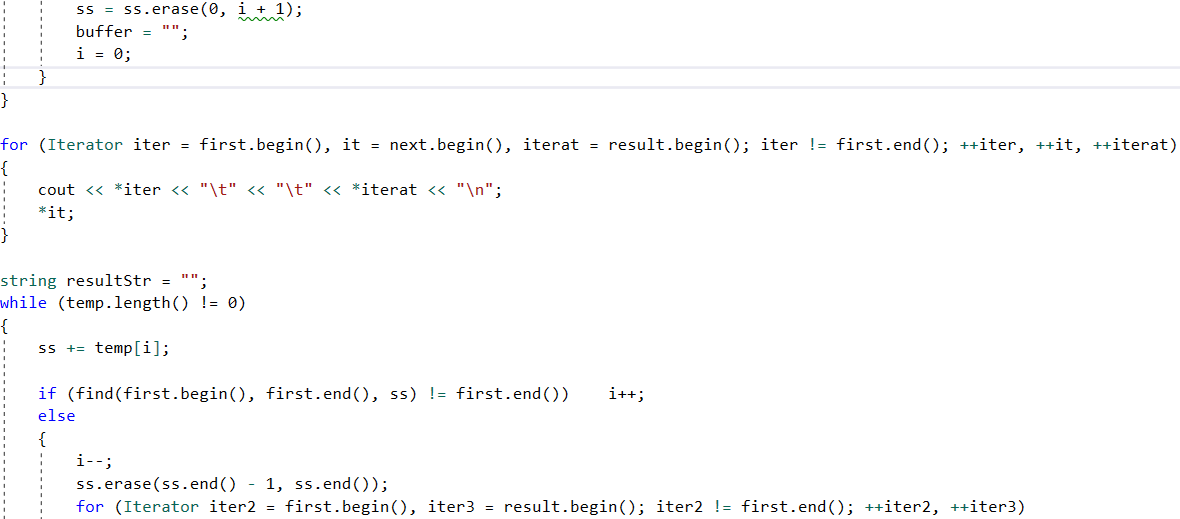
0001 0001 1000 0011 0100 0001 … и т.д. всего 76 битов. Если кодировать буквы по два бита, то потребуется 38 битов. Однако при увеличении размера исходной последовательности достигаемый по Лемпелю –Зива эффект будет значительным.

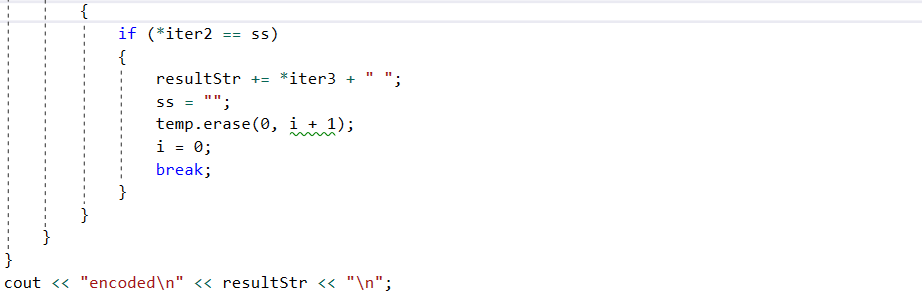
# **3 Задание**

1. Закодировать алгоритм на языке С (С++).
2. Продумать, как выполнить декомпрессию.









Дешифрование

