

**МИНОБРНАУКИ РОССИИ**  
**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ**  
**ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**  
**«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)**  
**Кафедра МО ЭВМ**

**КУРСОВАЯ РАБОТА**  
**по дисциплине «Алгоритмы и структуры данных»**  
**Тема: Статическое кодирование и декодирование файла методом Фано-**  
**Шеннона**

Студент гр. 0304

\_\_\_\_\_

Докучаев Р.А

Преподаватель

\_\_\_\_\_

Берленко Т.А.

Санкт-Петербург

2021

## ЗАДАНИЕ НА КУРСОВУЮ РАБОТУ

Студент Докучаев Р.А.

Группа 0304

Тема работы: Статическое кодирование и декодирование файла методом Фано-Шеннона. Вариант 1.

Исходные данные: На вход подаётся файл с закодированным или незакодированным содержимым.

Содержание пояснительной записки:

«Содержание», «Введение», «Реализация алгоритма», «Заключение».

Предполагаемый объем пояснительной записки:

Не 10 менее страниц.

Дата выдачи задания: 25.10.2021

Дата сдачи реферата: 24.12.2021

Дата защиты реферата: 24.12.2021

Студент

\_\_\_\_\_

Докучаев Р.А.

Преподаватель

\_\_\_\_\_

Берленко Т.А.

## СОДЕРЖАНИЕ

Содержание	3
1. Введение	4
1.1. Цель курсовой работы	4
1.2. Формулировка задания	4
1.3. Задачи	4
2. Реализация алгоритма	5
3. Демонстрация работы программы	7
Заключение	11
Список использованных источников	12

# **1. ВВЕДЕНИЕ**

## **1.1. Цель курсовой работы**

Была поставлена цель изучить алгоритм кодирования методом Фано-Шеннона и реализовать программу, способную кодировать или декодировать подающуюся на вход в виде файла информацию.

## **1.2. Формулировка задания**

"Демонстрация" – визуализация структур данных/алгоритмов. Демонстрация должна быть подробной и понятной (в том числе сопровождаться пояснениями), чтобы программу можно было использовать при объяснении используемой структуры данных и выполняемых с ней действий. Помимо демонстрации (визуализации) следует на каждом шаге выводить текстовые объяснения того, что происходит.

Статическое кодирование и декодирование текстового файла методом Фано-Шеннона – демонстрация

На вход подаётся файл с закодированным или незакодированным содержимым. Требуется раскодировать или закодировать содержимое файла методом Фано-Шеннона.

## **1.3. Задачи**

Для достижения поставленной цели требуется решить следующие задачи:

- Создание и разработка программного кода, который способен выполнять поставленные в исходных данных условия: кодирование и декодирование алгоритмом Фано-Шеннона данных, подаваемых в файле на вход
- Тестирование программы

## 2. РЕАЛИЗАЦИЯ АЛГОРИТМА

Все функции и классы для реализации алгоритма кодирования и декодирования Фано-Шеннона описаны и реализованы в файле sw.cpp.

Была создана структура Shannon Fano, имеющая поля data и length, а также оператор сравнения – равен ли один экземпляр структуры другому. Так как алгоритм Фано-Шеннона генерирует код переменной длины, то данные поля служат для того, чтобы хранить код и его длину.

Был создан класс-компаратор Symbol comp, который сравнивает пары символов и их весов, чтобы можно было создавать вектор символов по встречаемости. Это важная часть кодирования методом Фано-Шеннона, поскольку перед записью кодирующих символов необходимо отсортировать по невозрастанию символы исходного сообщения.

Была реализована функция std::ostream& operator<<, которая осуществляет вывод в поток кодов символа закодированного сообщения.

Были реализованы две основные функции, которые осуществляют основные действия при реализации кодирования методом Фано-Шеннона. std::map<char, Shannon Fano> Shannon Fano Table(std::string text) формирует вектор символов исходного сообщения и частоты их встречаемости в этом сообщении. Функция void Shannon Fano Table r(std::vector<std::pair<char, unsigned int>>& symbols, std::map<char, Shannon Fano>& table, int l, int r) рекурсивно обрабатывает исходный набор символов: делит на примерно равные по весу половины из поданных на вход символов и записывает 1 для большей половины, 0 для меньшей половины в качестве части кодировки символа.

Были реализованы две функции: void Shannon Fano encode(std::ifstream &file\_in, std::ofstream &file\_out) получает символы из исходного, записывает для каждого символа его биты в буфер в виде одного байта (unsigned char byte), когда этот байт заполняется — записывает его в файл и сбрасывает как сам байт, так и индекс текущего бита (size\_t bit) (обработка и кодирование происходят в функции Shannon Fano Table, по сути, данный метод отвечает за корректное считывание исходного сообщения и корректную запись полученной кодировки);

функция void Shannon\_Fano\_decode(std::ifstream &file\_in, std::ofstream &file\_out)  
тот же принцип работы, только происходит не кодирование исходного сообщения, а декодирование.

В головной функции main происходит открытие нужных файлов с исходными данными, открытие файлов для вывода информации и вызов функций Shannon\_Fano\_encode и Shannon\_Fano\_decode.

## ДЕМОНСТРАЦИЯ РАБОТЫ ПРОГРАММЫ

На следующих скриншотах терминала представлен процесс кодирования следующей фразы:

```
≡ Code
1 I love AaSD. Dokuchaev Roman
```

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	0
e	0
h	0
k	0
l	0
m	0
n	0
o	0
u	0
v	0

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	01
e	01
h	01
k	01
l	01
m	01
n	00
o	00
u	00
v	00

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	01
e	01
h	01
k	01
l	01
m	01
n	001
o	001
u	000
v	000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	01
e	01
h	01
k	01
l	01
m	01
n	001
o	001
u	0001
v	0000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	01
e	01
h	01
k	01
l	01
m	01
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	011
e	011
h	011
k	010
l	010
m	010
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	011
e	011
h	011
k	0101
l	0101
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	011
e	011
h	011
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	0111
e	0111
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:

.	1
A	1
D	1
I	1
R	1
S	1
a	1
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000



Таблица кодов Шеннона-Фано:	
.	11
A	11
D	10
I	10
R	10
S	10
a	10
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:	
.	11
A	11
D	101
I	101
R	101
S	100
a	100
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:	
.	11
A	11
D	101
I	101
R	101
S	1001
a	1000
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:	
.	11
A	11
D	1011
I	1011
R	1010
S	1001
a	1000
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:	
.	11
A	11
D	10111
I	10110
R	1010
S	1001
a	1000
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:	
.	111
A	110
D	10111
I	10110
R	1010
S	1001
a	1000
c	01111
e	01110
h	0110
k	01011
l	01010
m	0100
n	0011
o	0010
u	0001
v	0000

Таблица кодов Шеннона-Фано:	
.	111
A	1101
D	1100
I	10111
R	10110
S	1010
a	1001
c	1000
e	01111
h	01110
k	0110
l	01011
m	01010
n	0100
o	0011
u	0010
v	0001
	0000

≡ Decode  
1 I love AaSd. Dokuchaev Roman

Итоговая таблица кодов для данного исходного кода, а также содержание раскодированного бинарного файла, который был получен путём кодирования исходного, представлены выше

Второй пример аналогичен первому, только последовательность исходных символов проще.

Таблица кодов Шеннона-Фано:  
a | 1  
b | 1  
c | 0  
d | 0  
e | 0

Таблица кодов Шеннона-Фано:  
a | 1  
b | 1  
c | 01  
d | 00  
e | 00

Таблица кодов Шеннона-Фано:  
a | 1  
b | 1  
c | 01  
d | 001  
e | 000

Таблица кодов Шеннона-Фано:  
a | 11  
b | 10  
c | 01  
d | 001  
e | 000

≡ Code  
1 aaabbccddde

≡ Decode  
1 aaabbccddde

## **ЗАКЛЮЧЕНИЕ**

Была изучен алгоритм кодирования информации методом Фано-Шеннона, а также была реализована программа, реализующая данный алгоритм.

## **СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Алгоритм Фано-Шеннона/Хабр. // <https://habr.com/ru/post/137766/>
2. С. Рао "Освой самостоятельно C++ за 21 день"