Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования   
**Национальный исследовательский университет «МЭИ»**

**Лабораторная работа №3**

**«Кодирование информации методом Шеннона-Фано»**

Выполнил:

студент группы А-07-18

Востриков Р.В.

Вариант 8

Москва 2020

**1. Задание**

Блок №1  
Закодировать сообщения методом Шенона-Фано.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Буква | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Вероятность | 0,4 | 0,2 | 0,1 | 0,1 | 0,1 | 0,05 | 0,05 |

Блок №2

Написать программу для кодирования и декодирования с помощью функции XOR (исключающее ИЛИ).

**2. Описание результатов**

Блок №1

Имеется набор букв и соответствующих им вероятностей.

Устанавливаем для набора пустое начальное кодовое слово. Производим разбиение набора на две группы с примерно одинаковыми суммами вероятностей. К кодовому слову первой группы добавляем 0, к кодовому слову второй группы добавляем 1. Продолжаем процедуру до тех пор, пока у каждого сообщения не будет уникального кодового слова. По мере нахождения каждое кодовое слово добавляется в отдельный набор, где каждому кодовому слову соответствует исходное сообщение.

На вход подаётся сообщение, которое кодируется соответствующим кодовым словом из созданного набора и выводится.

Для наглядности построим таблицу:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Вероятность | Буква | 1-й бит | 2-й бит | 3-й бит | 4-й бит | 5-й бит | Кодовое слово |
| 0,4 | 1 | 0 | 0 | – | – | – | 00 |
| 0,2 | 2 | 1 | – | – | – | 01 |
| 0,1 | 3 | 1 | 0 | – | – | – | 10 |
| 0,1 | 4 | 1 | 0 | – | – | 110 |
| 0,1 | 5 | 1 | 0 | – | 1110 |
| 0,05 | 6 | 1 | 0 | 11110 |
| 0,05 | 7 | 1 | 11111 |

Средняя длина слова: L = 0,3\*2 + 0,2\*2 + 0,1\*2 + 0,1\*3 + 0,1\*4 + 0,05\*5 + 0,05\*5 = 2,4

При кодировании равномерным кодом кодовые слова выглядели бы так:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Буква | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Кодовое слово | 000 | 001 | 010 | 011 | 100 | 101 | 110 |

Средняя длина слова L = 3

Как можно заметить, средняя длина слова при кодировании методом Шеннона-Фано меньше, нежели при равномерном кодировании.

Блок №2

На вход программы поступает исходное сообщение и двоичная маска, по которой будет кодироваться сообщение. Каждый символ исходного сообщения кодируется с помощью операции XOR. Выводится результат. Это же закодированное сообщение кодируется еще раз через исходную маску. Тем самым получая декодированное сообщение.

**3. Код Программы (Блок №2)**

**Задание 1**

def aver\_codelength(codes,probs):

if len(codes) == len(probs):

aver = 0

for c, p in zip(codes,probs):

aver += len(c)\*p

return aver

else:

raise Exception("Количество кодовых слов не соответствует количеству вероятностей")

**Задание 2**

def is\_prefix(codes):

isprefix = True

length = len(codes)

i = 0

while(i < length and isprefix):

j = 0

while(j < length and isprefix):

if(i != j):

isprefix = not (codes[i].find(codes[j]) == 0)

j+=1

i+=1

return isprefix

**Задание 3**

def xor\_code(s, t):

key = it.islice(it.cycle(t), len(s)) # приведение длины ключа к длине сообщения

return bytes([a ^ b for a, b in zip(s, key)])

**4. Вывод**

Т.к. средняя длина слова, закодированного методом Шеннона-Фано, меньше средней длины слова, закодированного равномерным кодом, можно сделать вывод, что кодирование кодом Шеннона-Фано экономнее, чем кодирование равномерном кодом.

С помощью логической функции исключающее ИЛИ можно кодировать и декодировать сообщения однозначно.