**Лабораторная работа №1**

**«Изучение методов Шеннона – Фано и Хафмана по построению эффективных кодов»**

**1 Цель работы** - изучить возможности эффективного кодирования информации по методам Шеннона – Фано и Хаффмана.

**2 Основные теоретические сведения**

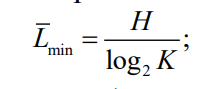
Кодирование, при котором обеспечивается минимальная средняя длина кодовых слов, называется ***эффективным*** (оптимальным). В эффективном коде символу, встречающемся чаще всего, присваивается наиболее короткая кодовая комбинация.

Задачи эффективного кодирования заключаются в следующем:

1) Запоминание максимального количества информации в ограниченной памяти.

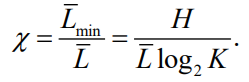
2) Обеспечение максимальной пропускной способности канала связи.

Эффективное кодирование базируется на теореме Шеннона о кодировании при отсутствии помех, согласно которой минимальная средняя длина кодовых слов определяется соотношением:

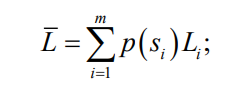


где H – энтропия источника сообщений, К – основание кода. Для двоичного кода, очевидно, что Lmin = H.

***Эффективностью*** кода х называется отношение Lmin к реально достигнутой в данном коде средней длине кодовых слов :



Средняя длина кодовых комбинаций может быть найдена следующим образом:



где Li - длина кодовой комбинации, соответствующей символу si из алфавита размером ***m*** .

Для случая отсутствия статистической взаимосвязи между символами метод построения эффективных кодов впервые был предложен Шенноном и Фано.

Для двоичного кода метод Шеннона-Фано сводится к следующему:

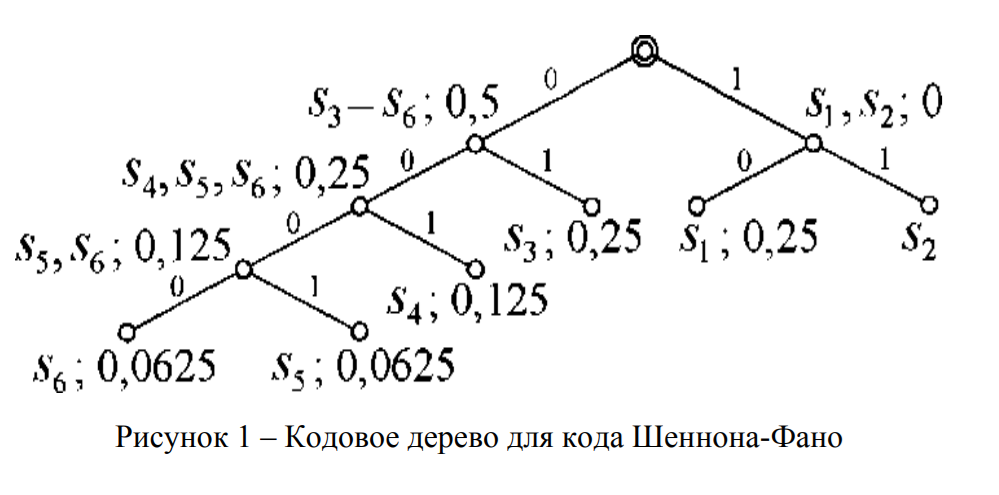
1) Буквы алфавита располагаются в порядке убывания вероятностей.

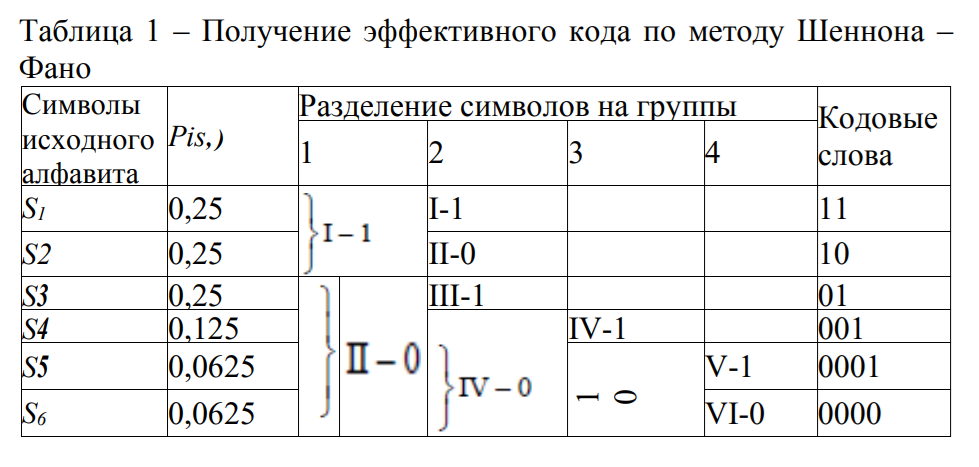
2) Алфавит букв разбивается на две группы таким образом, чтобы суммарные вероятности букв обеих групп были по возможности равны. Первой группе присваивается символ 1, второй символ – 0.

3) Каждую из образованных групп вновь делят на две части с приблизительно равными суммарными вероятностями и присваивают им 1 и 0. Таким образом, получают вторые цифры кода.

4) Процесс повторяется до тех пор, пока в каждой подгруппе не останется по одной букве.

Рассмотрим пример построения кода Шеннона-Фано для алфавита из шести символов (таблица 1). Кодовое дерево, соответствующее полученному коду Шеннона-Фано, представлено на рисунке 1.





Метод Шеннона-Фано не всегда приводит к однозначному построению кода. От указанного недостатка свободен метод Хаффмана.

Для двоичного кода метод Хаффмана сводится к следующему:

1) Буквы алфавита выписываются в столбец в порядке убывания вероятностей.

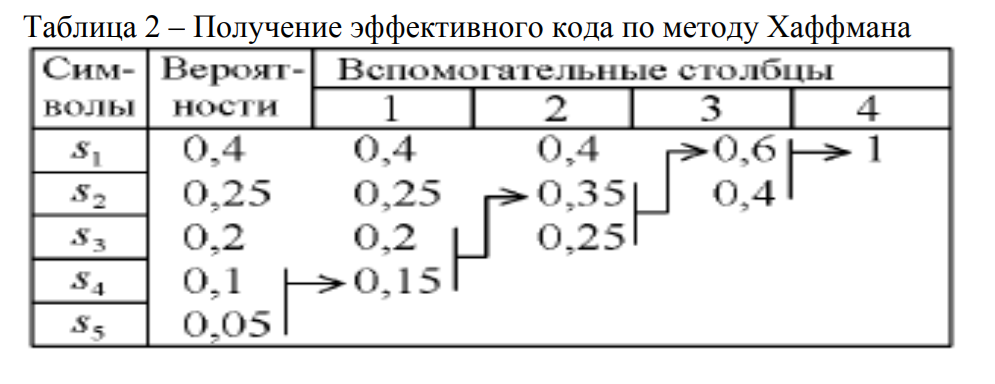
2) Две последние буквы объединяются в одну вспомогательную букву, которой приписывается суммарная вероятность.

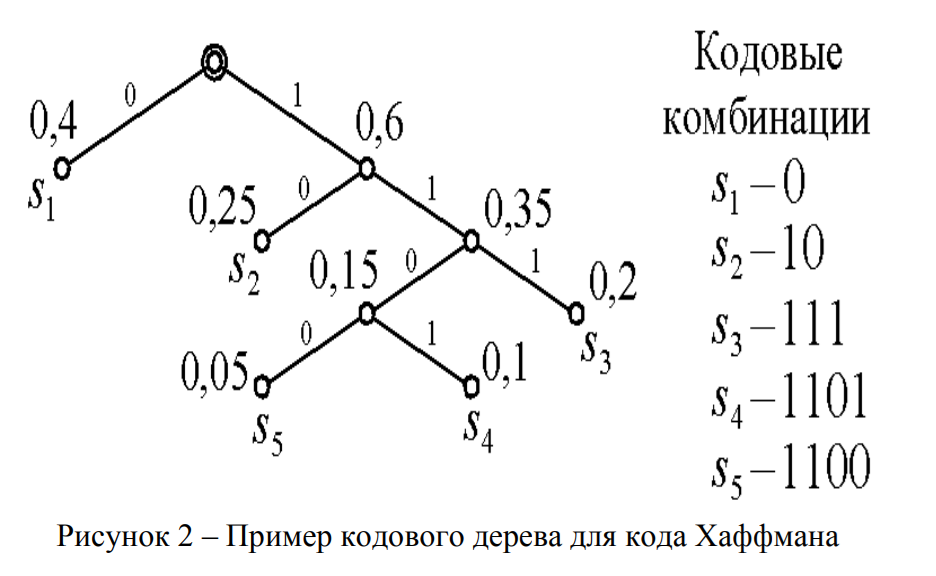
3) Вероятности букв, участвующих в объединении и полученная суммарная вероятность вновь располагаются в порядке убывания вероятностей в дополнительном столбце, а две последние буквы объединяются.

4) Процесс продолжается до тех пор, пока не будет получена единственная вспомогательная буква с суммарной вероятностью, равной 1.

Для получения кодовой комбинации, соответствующей данной букве необходимо проследить путь перехода по строкам и столбцам таблицы.

Рассмотрим пример построения кода Хаффмана для алфавита из восьми символов (таблица 2). Результат показан на рисунке 2 в виде кодового дерева соответствующего коду Хаффмана.





**3 Порядок выполнения работы**

Данная лабораторная работа предполагает выполнение следующих этапов:

1) Изучить методические указания к лабораторной работе.

2) Пройти собеседование с преподавателем и получить задание для выполнения работы.

3) На основе заданного первичного алфавита и вероятностей появления символов этого алфавита (табл. 3.3) получить в форме таблицы двоичный код Шеннона-Фано.

5) Определить эффективность кода, полученного по методу Шеннона-Фано.

6) Выполнить пункты 3, 4, 5 для метода Хаффмана.

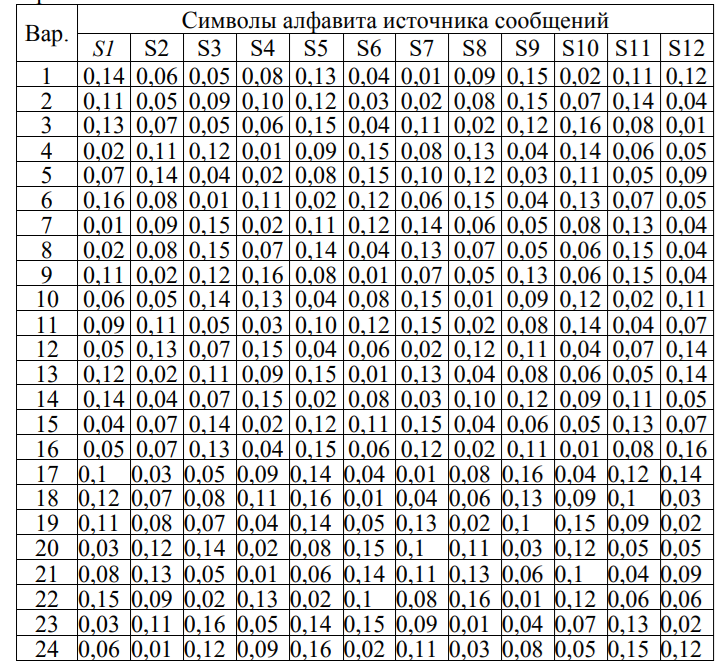
7) Сравнить эффективности методов Шеннона-Фано и Хаффмана. Сделать выводы о полученных результатах.

8) Оформить и защитить отчет по выполнению лабораторной работы.

**4 Варианты заданий**

Номер варианта студента определяется как порядковый номер в журнале преподавателя.

Таблица 3 – Вероятности появления символов для различных вариантов



**5 Контрольные вопросы**

1. В чем заключается сущность эффективного кодирования?

2. Каковы основные задачи эффективного кодирования?

3. Как определяется средняя длина кодового слова?

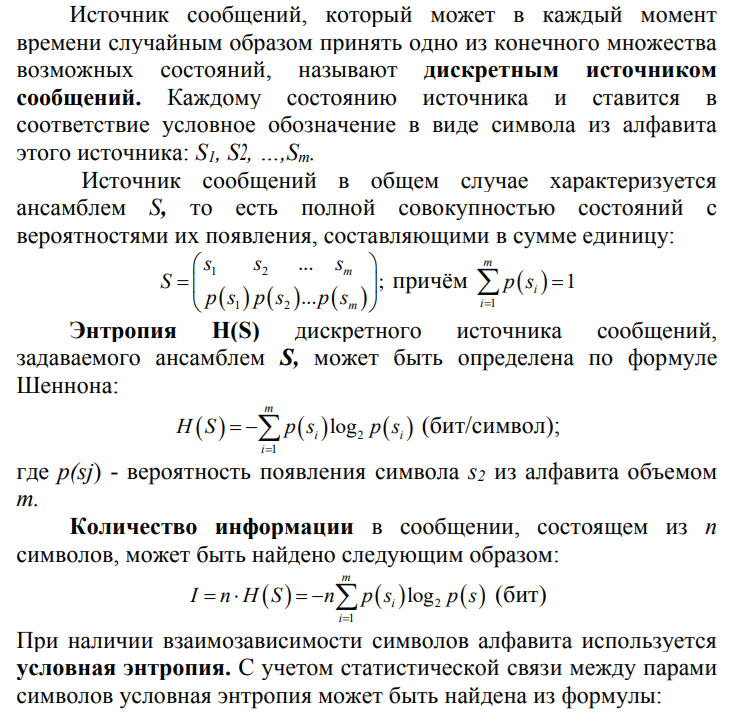
4. Чему равна нижняя граница эффективного кодирования?

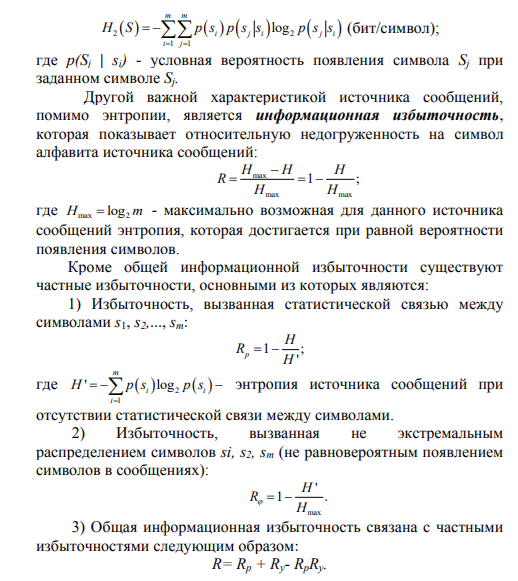
5. Как определяется эффективность кода?

6. В каком случае метод Шеннона-Фано гарантированно обеспечивает получение эффективного кода?

7. С помощью какой операции в методе Хаффмана обеспечивается получение вспомогательных символов?

**6 Дополнительные сведения**



****