МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное образование учреждение высшего образования

«Курганский государственный университет»

Кафедра «Программное обеспечение автоматизированных систем»

РФ КГУ 09.03.04. КР24.360091

**КОНТРОЛЬНАЯ РАБОТА**

**по дисциплине**

**«Теория информации»**

Выполнил студент гр. ИТ-0940323\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ / Белоусов К.А /

Проверил канд. техн. наук, доцент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ / Семахин А.М. /

# **СОДЕРЖАНИЕ**

[СОДЕРЖАНИЕ 2](#_Toc184719127)

[ВВЕДЕНИЕ 3](#_Toc184719128)

[1. МЕТОДЫ ЭФФЕКТИВНОГО КОДИРОВАНИЯ 4](#_Toc184719129)

[1.1 Метод Шеннона-Фано 4](#_Toc184719130)

[1.2 Метод Хаффмана 5](#_Toc184719131)

[2. ВЫПОЛНЕНИЕ КОНТРОЛЬНОЙ РАБОТЫ 6](#_Toc184719132)

[2.1 Построение кодовой таблицы 6](#_Toc184719133)

[2.2 Настройка кодера 9](#_Toc184719134)

[2.3 Настройка декодера 11](#_Toc184719135)

[2.4 Восстановление закодированного сообщения 12](#_Toc184719136)

[ЗАКЛЮЧЕНИЕ 14](#_Toc184719137)

# **ВВЕДЕНИЕ**

Теория информации — это раздел математики, исследующий процесс хранения, преобразования и передачи информации. Она занимается измерением количества информации, изучением её свойств и установлением ограничений для эффективных систем передачи данных. Для решения этих задач теория информации использует методы теории вероятностей и математической статистики, а также различные математические модели.

Задачи теории информации:

1. Исследование процесса кодирования сообщения.
2. Анализ передачи закодированного сообщения через канал
3. Исследование восстановления исходного сообщения

**Постановка задачи:** выполнить кодирование сообщения методами эффективного кодирования, настроить кодер и декодер для передачи кодовой комбинации от источника к приёмнику.

# **МЕТОДЫ ЭФФЕКТИВНОГО КОДИРОВАНИЯ**

## **Метод Шеннона-Фано**

Учитывая статистические свойства источника сообщений, можно минимизировать среднее число двоичных символов, требующихся для выражения одной буквы сообщений, что при отсутствии шума позволяет уменьшить время передачи или емкость запоминающего устройства. Такое эффективное кодирование базируется на основной теореме Шеннона для каналов без шума.

К. Шеннон доказал, что сообщения, составленные из букв некоторого алфавита, можно закодировать так, что среднее число двоичных символов на букву будет приближаться к энтропии источника этих сообщений, но не менее этой величины.

Теорема не указывает конкретный способ кодирования, однако из нее следует, что для эффективного кодирования каждый символ должен нести максимальную информацию. Следовательно, каждый символ должен принимать значения 0 или 1 по возможности с равными вероятностями, и каждый выбор должен быть независим от значений предыдущих символов.

Методы построения эффективных кодов были предложены Шенноном и Фано, и их алгоритмы во многом схожи, поэтому получившийся код называется кодом Шеннона-Фано.

Процесс кодирования по методу Шеннона-Фано следующий:

* 1. Алфавит сообщений сортируется в порядке убывания вероятностей появления символов.
  2. Алфавит делится на две группы таким образом, чтобы суммы вероятностей для каждой группы были максимально равны.
  3. Символы первой группы получают в качестве первого бита 1, а символы второй группы — 0.
  4. Каждая из полученных групп снова делится на две подгруппы с равными суммами вероятностей и процесс повторяется.
  5. Процесс продолжается до тех пор, пока в каждой подгруппе не останется только один символ.

В результате получается префиксный код, который гарантирует, что ни один код не будет являться префиксом другого. Это свойство важное для однозначности декодирования и позволяет записывать текст без разделителей.

Рассмотрим алфавит из 8 букв. Ясно, что при обычном (не учитывающем вероятностей встречаемости их в сообщениях) кодировании для представления каждой буквы требуется 3 символа (), где *M* – количество букв в алфавите.

Наибольший эффект сжатия достигается, когда вероятности символов представляют собой отрицательные целочисленные степени двойки. В этом случае среднее количество бит на символ точно равно энтропии. В более общем случае, для алфавита из 8 символов, среднее количество бит на символ будет меньше 3, но больше энтропии алфавита *H(M)*.



## **Метод Хаффмана**

Он гарантирует однозначное построение кода с наименьшим для данного распределения вероятностей средним числом символов на букву. Для двоичного кода методика сводится к следующему. Буквы алфавита сообщений выписываются в первый столбец в порядке убывания вероятностей. Две последние вероятности объединяются в одну вспомогательную, которой приписывается суммарная вероятность. Вероятности букв, не участвующих в объединении, и полученная суммарная вероятность снова располагаются в порядке убывания вероятностей в дополнительном столбце, а две последние вероятности снова объединяются. Процесс продолжается до тех пор, пока не получим единственную вспомогательную вероятность, равную единице.

Алгоритм метода Хаффмана включает этапы:

1. Подсчет частот символов в исходном тексте
2. Создание списка узлов, где каждый символ - отдельный узел с его частотой
3. Поиск двух узлов с наименьшими частотами и объединение их в новый узел с суммарной частотой
4. Присвоение битов 0 и 1 ветвям нового узла (левой - 0, правой - 1)
5. Повторение шагов 3-4 пока все узлы не будут объединены в одно дерево
6. Определение кодов символов путем прохода от корня до листьев, записывая последовательности битов

# **ВЫПОЛНЕНИЕ КОНТРОЛЬНОЙ РАБОТЫ**

## **Построение кодовой таблицы**

Для осуществления кодирования исходного сообщения с использованием эффективного кодирования Хаффмана было разработано визуальное программное приложение на языке C++

Вводим исходный текст (Рисунок 1.)

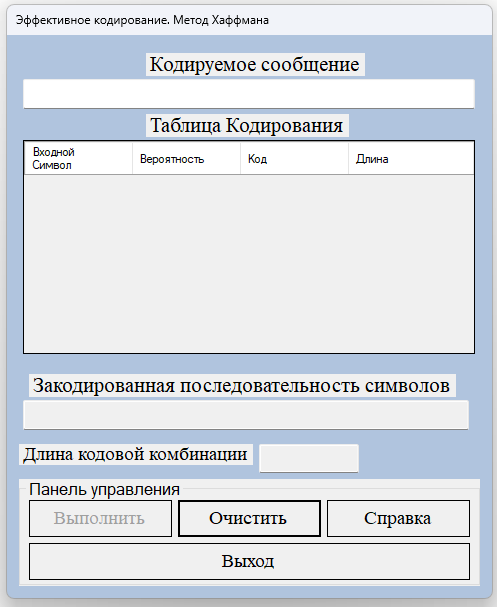


Рисунок 1. – Вид программы на этапе ввода кодируемого сообщения.

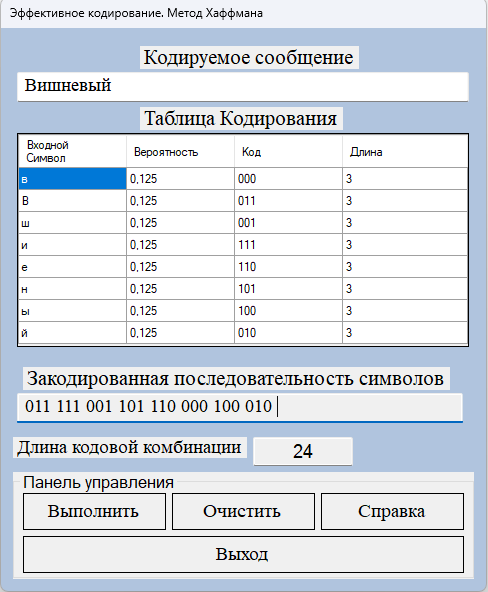
На экране выводятся вероятность появления и коды символов (Рисунок 2.)

Рисунок 2. – Результат кодирования исходного сообщения

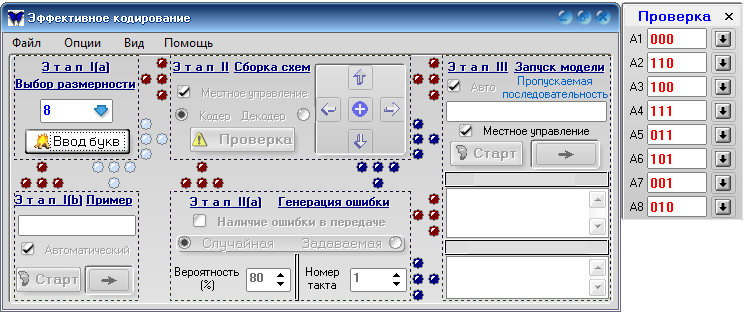
Подставляем полученные коды, предварительно перевернув их в программу Effectcoding.exe. (Рисунок 3)

Рисунок 3. – Основная панель приложения Effectcoding

## **Настройка кодера**

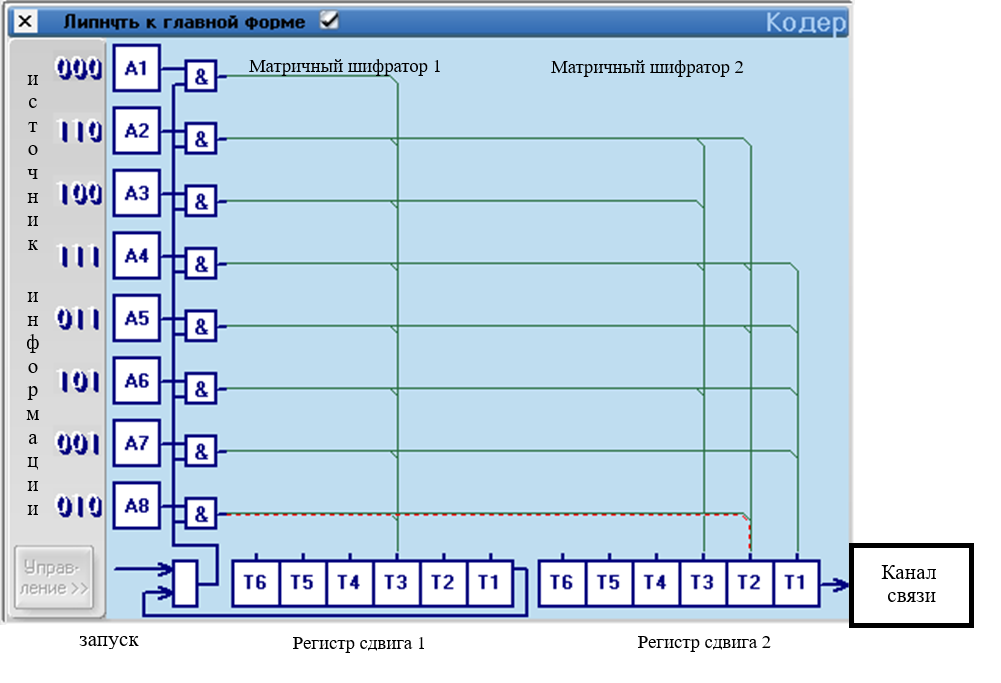
Собираем кодер. (Рисунок 4.)

Рисунок 4. – Настроенный кодер

Устройство включает схему получения моментов считывания очередной буквы (матричный шифратор 1 с регистром сдвига 1) и шифратор букв (матричный шифратор 2 с регистром сдвига 2). Число горизонтальных шин шифраторов равно числу кодируемых букв, а число вертикальных шин в каждом из них равно числу символов в самой длинной комбинации используемого кода. Схема рассчитана на использование алфавита из 6-8 букв (сообщений). Источник информации удовлетворяет требованиям идеального источника по Хартли, т.е. в каждый данный момент он возбуждает шину только одной буквы, выставляя на нее сигнал "1". В регистре 1 появляется сигнал "1" в разряде соответствующем длине кодовой комбинации буквы, что обеспечивается диодным переходом. В регистре 2 появляется код соответствующий букве, шина которой возбуждена. Что тоже обеспечивается соответствующими диодными переходами с возбужденной шины на триггеры регистра 2. Сдвигающими импульсами код буквы последовательно выталкивается в канал связи, а единица "вытолкнутая" из регистра 1 разрешает источнику выдать очередную букву.

## **Настройка декодера**

1. Собираем декодер (Рисунок 5.)

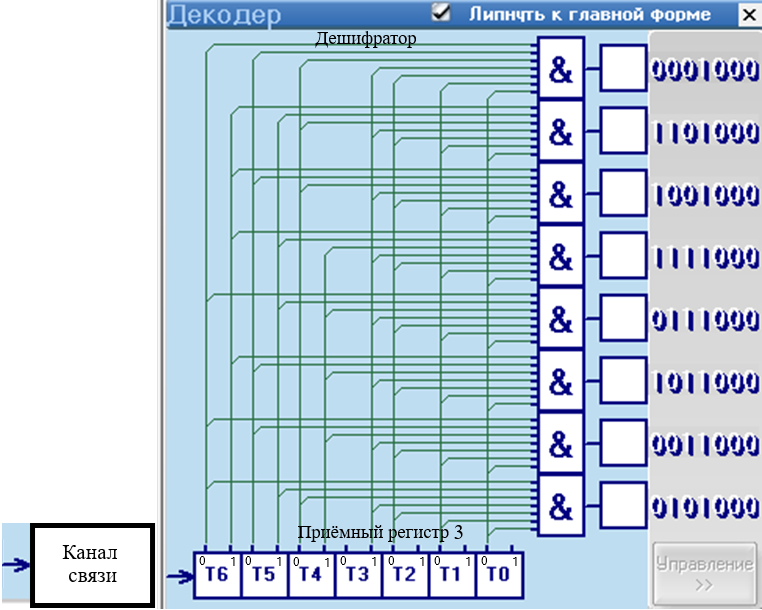


Рисунок 5. – Настроенный декодер

Информация поступает из канала связи в приемный регистр 3. Регистр 3 состоит из триггеров, количество которых на один больше, чем самая длинная кодовая комбинация одной из букв алфавита. В начальном состоянии в регистре 3 в триггере T6 записана “1”. Перемещаясь по регистру сдвигающими импульсами по мере приема информации из линии связи, она информирует о приеме очередной буквы. Когда кодовая комбинация буквы разместится в регистре 3 полностью, в этот момент должна схема совпадения “И”, соответствующая принятой букве.

Для этого на все входы схемы "И" этой буквы должны быть поданы сигналы "1" через диоды, подключенные к триггерам регистра 3, состояния которых соответствуют кодовой комбинации принятой буквы. Если триггер находится в состоянии "1”, то диод подключается к правому или единичному выходу, а если в состоянии "0", то к левому или нулевому выходу. Срабатывание схемы "И" принятой буквы индицирует ее и через схему “ИЛИ” сбрасывает в ноль все разряды регистра 3, кроме первого, в котором снова выставляется "1". После этого схема декодера готова к приему очередной буквы.

## **Восстановление закодированного сообщения**

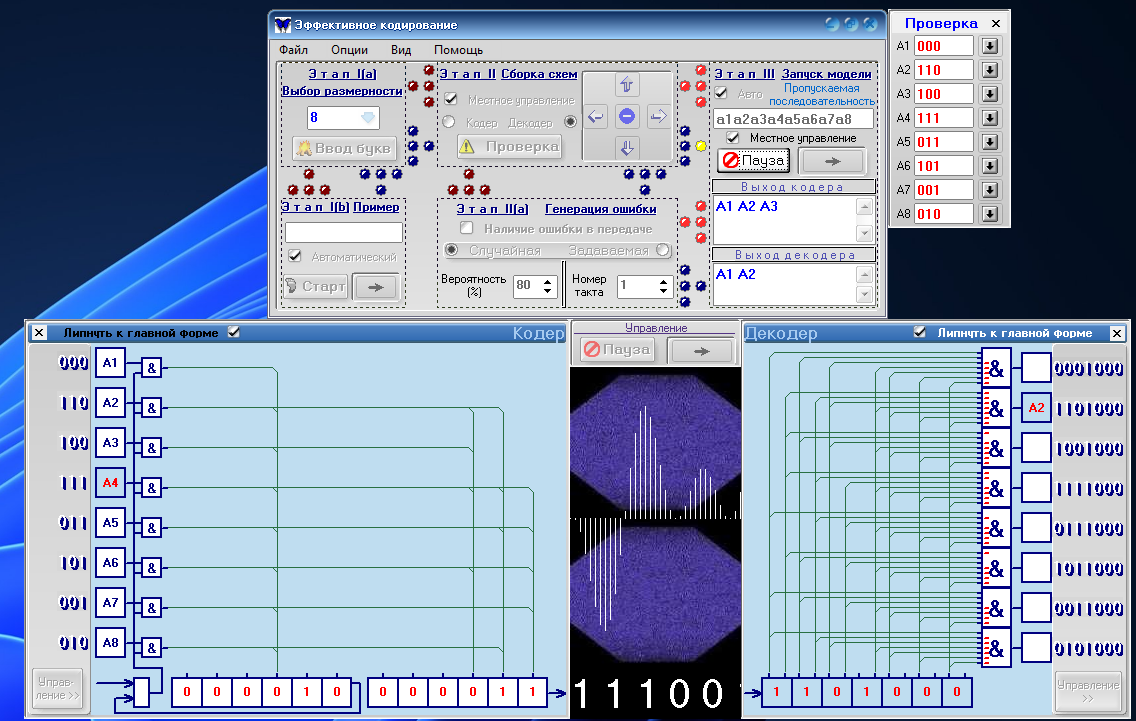
Запускаем последовательность a1a2a3a4a5a6a7a8 (рисунок 6.)

Рисунок 6. – Результат запуска последовательности

На кодере загораются слова, которые в данный момент кодируются и передаются в канал связи (со сдвигом по регистрам 1 и 2), на декодере с тактовой задержкой равной 6 тактам происходит распознавание сигнала/кода (со сдвигом по регистру 3) {использование в качестве слов не префиксных кодов приведёт к неверному декодированию}: красные штрихи у сумматоров означают соответствие сигнала с триггера и схемы, их отсутствие означает обратное; цветом выделяются сумматоры, в которых распознано слово, и триггеры, нашедшие хотя бы одно вышеуказанное соответствие.

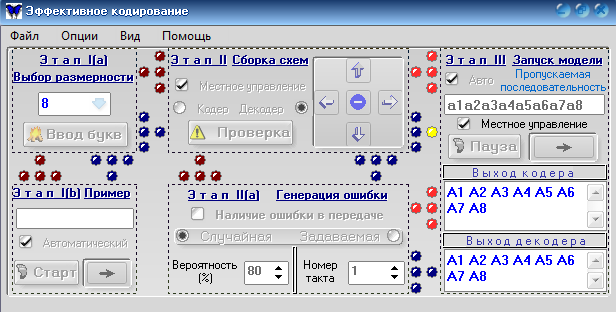
При восстановлении закодированного сообщения, оно было полностью восстановлено (рисунок 7):

Рисунок 7 – Положительный результат восстановления сообщения

# **ЗАКЛЮЧЕНИЕ**

В процессе выполнения контрольной работы было создано визуальное приложение, осуществляющее кодирование исходных данных с применением метода эффективного кодирования Хаффмана. Практическая реализация включала разработку интерфейса, настройку процесса кодирования и декодирования.

В ходе работы были изучены основные принципы эффективного кодирования, получены навыки реализации алгоритма Хаффмана, а также расширены знания в области обработки и передачи информации.

СПИСОК ЛИТЕРАТУРЫ

1. Семахин А. М. Теория информации: методические указания к выполнению практических и контрольных работ для студентов направлений подготовки 09.03.04, 09.03.03 – Курган: Издательство Курганского государственного университета, 2022 –45 с.
2. Семахин А. М. Теория информации: учебное пособие – Курган: Издательство Курганского государственного университета, 2023 –164 с.
3. Построение и реализация эффективных кодов // Pandia. – URL: https://pandia.ru/text/79/295/29822.php (дата обращения: 10.12.2024).