Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информационных технологий

Кафедра информационных систем и технологий

Специальность 1-40 01 01 Программное обеспечение информационных технологий

**Отчёт по лабораторной работе №1**

**«Элементы теории информации. Параметры и характеристики дискретных информационных систем»**

Исполнитель:

Студент 3 курса группы 4

Гурина К. С.

Руководитель:

Ассистент Сазонова Д. В.

1. **Цель и задачи работы**

Цель: приобретение практических навыков расчета и анализа параметров и информативных характеристик дискретных ИС.

Задачи:

1. Закрепить теоретические знания по основам теории информации.

2. Разработать приложение для расчета и анализа параметров и информативных характеристик дискретных ИС.

3. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**2. Теоретические сведения**

Передача информации (данных) осуществляется между двумя абонентами, называемыми источником сообщения (ИcС) и получателем сообщения (ПС). Третьим элементом информационной системы является канал (среда) передачи, связывающий ИсС и ПС.

Отображение сообщения обеспечивается изменением какой-либо физической величины, характеризующей процесс (например, амплитуда, частота, фаза). Эта величина является *информационным* *параметром сигнала* (в общем случае–информационной системы).

Сигналы, как и сообщения, могут быть непрерывными и дискретными. Информационный параметр непрерывного сигнала с течением времени может принимать любые мгновенные значения в определенных пределах. Непрерывный сигнал часто называют аналоговым, а каналы и устройства, функционирующие на основе такого типа сигналов – аналоговыми. Дискретный сигнал (устройство или канал передачи) характеризуется конечным числом значений информационного параметра.

Дискретные сообщения состоят из последовательности дискретных знаков. Часто этот параметр принимает всего два значения (0 или 1).

Построение сигнала по определенным правилам, обеспечивающим соответствие между сообщением и сигналом, называют кодированием. Кодирование в широком смысле – преобразование сообщения в сигнал. Кодирование в узком смысле – представление исходных знаков, называемых символами, в другом алфавите с меньшим числом знаков.

Важнейшая характеристика источника, получателя или канала – алфавит.

Алфавит, А – это общее число знаков или символов (N), используемых для генерации или передачи сообщений. Характеристикой алфавита является его мощность, N – количество символов, на основе которых формируется сообщение

Информационной характеристикой алфавита (источника сообщений на основе этого алфавита) является энтропия. Этот термин применительно к техническим системам был введен К. Шенноном и Р. Хартли.

Энтропию алфавита А{} по Шеннону рассчитывают по следующей формуле:

(1.1)

где P() – вероятность P(ξ=); – элемент алфавита,.

С физической точки зрения энтропия показывает, какое количество информации приходится в среднем на один символ алфавита.

Частным случаем энтропии Шеннона является энтропия Хартли. Дополнительным условием при этом является то, что все вероятности одинаковы и постоянны для всех символов алфавита.

С учетом этого формулу (1.1) можно преобразовать к виду:

(1.2)

Сообщение Хk, которое состоит из k символов, должно характеризоваться определенным количеством информации I(Хk):

(1.3)

Здесь Н(А)– энтропия алфавита с соответствующим распределением вероятностей р().

**3. Практическая часть**

**Практическое задание:**

Создать приложение для расчета и анализа параметров и информативных характеристик дискретных ИС, с помощью которого:

а) рассчитать энтропию указанного преподавателем алфавитов: один – на латинице, другой – на кириллице (по формуле (2.1) – перейти от частоты появления каждого символа алфавита к соответствующей вероятности); в качестве входного может быть принят произвольный электронный текстовый документ на основе соответствующего алфавита; частоты появления символов алфавитов оформить в виде гистограмм (можно воспользоваться приложением MSExcel);

б) для входных документов, представленных в бинарных кодах, определить энтропию бинарного алфавита;

в) используя значения энтропии алфавитов, полученных в пп. а) и б), подсчитать количество информации в сообщении, состоящем из собственных фамилии, имени о отчества (на основе исходного алфавита – а) и в кодах ASCII–б); объяснить полученный результат;

г) выполнить задание п. в) при условии, что вероятность ошибочной передачи единичного бита сообщения составляет: 0.1; 0.5; 1.0.

**Ход работы**

Для выполнения первого и второго задания было выбрано три алфавита: белорусский, английский и бинарный. Код этих алфавитов представлен на рисунке 3.1.

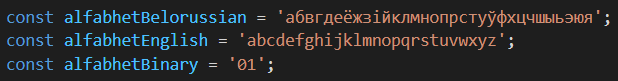


Рисунок 3.1 – Алфавиты

Для расчета энтропии алфавитов была разработана функция GetEntropyShanon, которая рассчитывает энтропию по формуле Шеннона: . Код этой функции представлен на рисунке 3.2.

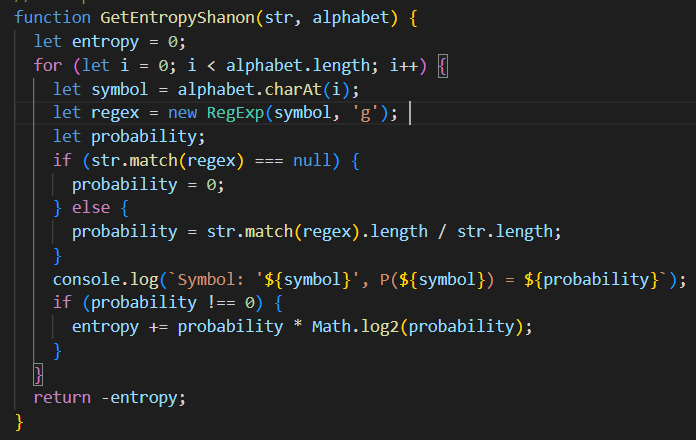


Рисунок 3.2 – Функция для расчёта энтропии по Шеннону

Частным случаем энтропии Шеннона считается энтропия Хартли. Для расчета энтропии алфавитов была разработана функция GetEntropyHartley, которая рассчитывает энтропию по формуле Хартли: . Код этой функции представлен на рисунке 3.3.

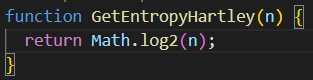


Рисунок 3.3 – Функция для расчёта энтропии по Хартли

В результате получаем такие результаты расчёта энтропии белорусского и английского алфавитов для исходных текстов, которые представлены на рисунке 3.4

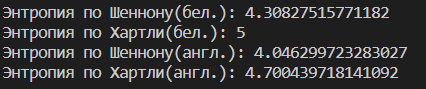


Рисунок 3.4 – Энтропия белорусского и английского алфавитов

На основании частот появления символов, рассчитанных в предыдущем шаге, созданы гистограммы. Данные гистограммы отображены на рисунках 3.5 и 3.6.



Рисунок 3.5 – Гистограмма частот появления символов для белорусского алфавита



Рисунок 3.6 – Гистограмма частот появления символов для английского алфавита

Для расчёта энтропии для бинарного алфавита нужно было перевести текст в бинарный текст. Для этого была разработана функция convbinary, код которой представлен на рисунке 3.7.

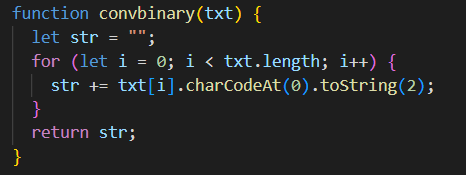


Рисунок 3.7 – Функция для перевода текста в бинарное представление

Результат использования функции GetEntropyShanon вместе с convbinary представлен на рисунке 3.8.



Рисунок 3.8 – Энтропия для бинарного алфавита

Следующим шагом нужно было рассчитать количество информации в сообщении, состоящем из собственных фамилии, имени о отчества. Для этого была использована формула *I*(*Xk*) *= H*(*A*)*⋅ k*. Код представлен на рисунке 3.9.

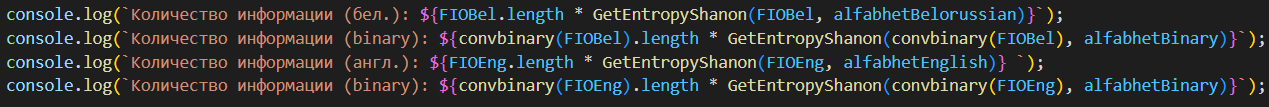


Рисунок 3.9 – Расчет количества информации

В результате получаем следующий вывод:

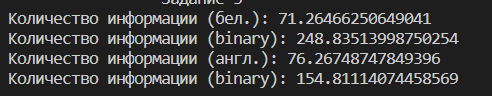


Рисунок 3.10 – Расчёт количества информации в сообщении

Следующим шагом нужно было подсчитать количество информации для бинарного алфавита при условии, что вероятность ошибочной передачи единичного бита сообщения составляет 0.1, 0.5, 1.0.

Поскольку вероятность ошибки отлична от нуля, то переданное сообщение может содержать ошибки. Количество информации в таком сообщении будет определяться не энтропией двоичного алфавита, а эффективной энтропией алфавита или пропускной способностью канала, которую можно рассчитать по следующей формуле: , где .

В результате получим следующий вывод:

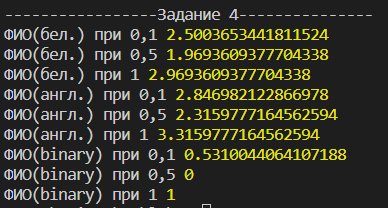


Рисунок 3.11 – Вывод терминала для бинарного алфавита

**Вывод**

В ходе лабораторной работы были приобретены практические навыки расчёта и анализа параметров и информативных характеристик дискретных ИС, закреплены теоретические знания по основам теории информации.

Также было разработано приложение для расчета и анализа параметров и информативных характеристик дискретных ИС.