Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

**Информационная безопасность**

Студент: Козека Е. М.

ФИТ 3 курс 4 группа

Преподаватель: Нистюк О. А.

Минск 2025

**Лабораторная работа №2. Элементы теории информации. Параметры и характеристики. Дискретных информационных систем**

**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.

**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.

Отметим также, что и в системах с хранением информации всегда можно выделить ИcС и ПС. В данном случае каналом передачи здесь выступает устройство хранения информации (память). Например, при записи данных в ОЗУ (оперативное запоминающее устройство) компьютера в качестве ИcС и ПС может выступать процессор (соответственно при записи и чтении данных).

Таким образом, простейшая информационная система состоит из трех элементов: источника сообщения, канала передачи сообщения и получателя сообщения.

**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). Сообщение или канал его передачи на основе этих двух значений сигнала называют двоичным или бинарным.

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. Оно осуществляется с целью повышения надежности и преобразования сигналов к виду, удобному для передачи по каналам связи.

Важнейшая характеристика источника, получателя или канала – **алфавит**. Алфавит, А – это общее число знаков или символов (N), используемых для генерации или передачи сообщений. Символы алфавита будем обозначать через {аi}, где 1 ≤ i ≤ N; N – мощность алфавита.

Информационной характеристикой алфавита (источника сообщений на основе этого алфавита) является **энтропия**.

Этот термин применительно к техническим системам был введен К. Шенноном и Р. Хартли.

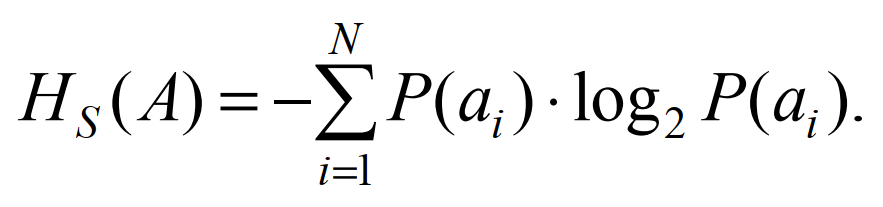
С физической точки зрения энтропия алфавита показывает, какое количество информации приходится в среднем на один символ алфавита.

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

**Ход работы**

**Задание 1.** Рассчитать энтропию указанных 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. Для входных документов, представленных в бинарных кодах, определить энтропию бинарного алфавита

Энтропию алфавита по К. Шеннону рассчитывают по следующей формуле:



Для выполнения задания были использованы турецкий и русский алфавиты. Для расчёта использовался следующий код на языке JavaScript:

const calculateSymbolsFrequency = (file, alphabet) => {

    let contents = fs.readFileSync(file, { encoding: "utf-8" });

    let resultFrequency = {};

    let symbolsCount = 0;

    if (alphabet === BINARY\_ALPHABET) {

        contents = convertStringToBinary(contents);

    }

    else {

        contents = contents.toLowerCase();

    }

    for (let symbol of contents) {

        if (alphabet.includes(symbol)) {

            resultFrequency[symbol] = (resultFrequency[symbol] || 0) + 1;

            symbolsCount++;

        }

    }

    for (let key in resultFrequency) {

        resultFrequency[key] = resultFrequency[key] / symbolsCount;

    }

    return { symbolsCount, resultFrequency };

};

const exportHistogram = async (file, frequency, alphabet) => {

    const opts = {

        chart: "column",

        titles: ["Частота"],

        fields: alphabet,

        data: { "Частота": frequency },

        chartTitle: "Частота появления символов в алфавите"

    };

    return new Promise((resolve, reject) => {

        xlsxChart.generate(opts, (err, data) => {

            if (err) {

                reject(err);

            } else {

                fs.writeFileSync(path.join(\_\_dirname, "charts", file), data);

                console.log(`Гистограмма сохранена: ${file}`);

                resolve();

            }

        });

    });

};

const convertStringToBinary = (str) => {

    return str.split("").map((char) => char.charCodeAt(0).toString(2)).join("");

};

const calculateEntropy = (frequency) => {

    let entropy = 0;

    for (let key in frequency) {

        entropy += frequency[key] \* Math.log2(frequency[key]);

    }

    return -entropy;

};

Листинг 2.1 – Функции для задания 1

Гистограмма частоты появления символов в тексте на турецком языке:

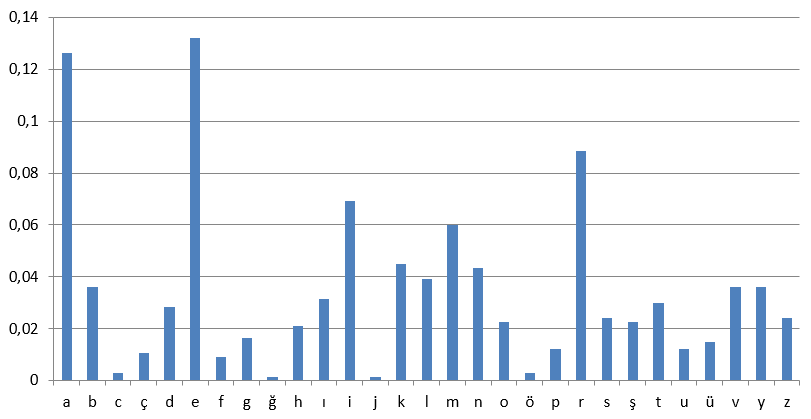


Рисунок 2.1 – Гистограмма для латиницы

Гистограмма частоты появления символов в тексте на русском языке:

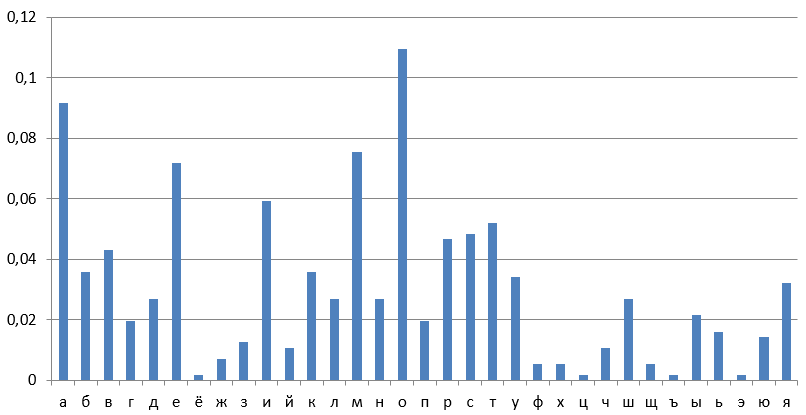


Рисунок 2.2 – Гистограмма для кириллицы

Результат работы приложения:

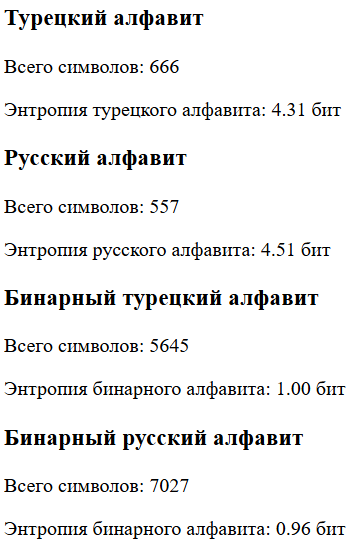


Рисунок 2.3 – Работа функций для расчета энтропии

**Задание 2.** Используя полученные значения энтропии алфавитов подсчитать количество информации в сообщении, состоящем из собственных имени, фамилии и отчества.

Дополним разработанный в предыдущем задании код следующим методом:

const calculateInformationCount = (message, alphabetEntropy) => {

    return message.length \* alphabetEntropy;

};

Листинг 2.2 – Функция для задания 2

Результат работы приложения:

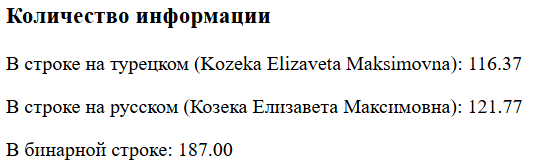


Рисунок 2.4 – Работа функции для расчета количества информации

**Задание 3.** Выполнить задание 2 при условии, что вероятность ошибочной передачи единичного бита сообщения составляет: 0.1, 0.5, 1.

Дополним разработанный в предыдущем задании код следующим методом:

const calculateEffectiveEntropy = (entropy, p) => {

    const q = 1 - p;

    let conditionalEntropy = -p \* Math.log2(p) - q \* Math.log2(q);

    return entropy - (Number.isNaN(conditionalEntropy) ? 0 : conditionalEntropy);

};

Листинг 2.3 – Функция для задания 4

Результат работы приложения:

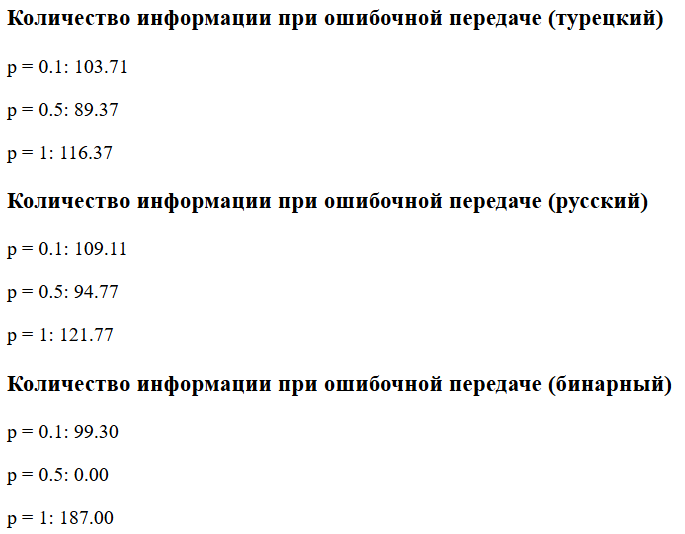


Рисунок 2.5 – Расчет количества информации при наличии ошибок передачи

**Вывод:** в ходе выполнения данной лабораторной работы была определена энтропия турецкого, русского и бинарного алфавитов, а также было определено количество информации в сообщениях на основе этих алфавитов. Кроме того, было определено количество информации при условии, что есть вероятность ошибочной передачи единичного бита.