Учреждение образования

«Белорусский государственный технологический университет»

**Лабораторная работа №1**

Расчет и анализ параметров и информативных характеристик дискретных ИС

Выполнил:

Студент 3 курса 5 группы ФИТ

Коршун Никита Игоревич

2024

**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 (можно воспользоваться приложением MS Excel);

Для выполнения задания были выбраны 2 языка: кириллица – македонский, латиница – турецкий. Было создано приложение используя язык разработки С++, для вычисления энтропии была создан отдельный метод. Она суммирует вероятности каждого символа, умноженные на логарифм по основанию 2 от этих вероятностей, и возвращает отрицательное значение этой суммы. Для вычисления была использована формула К. Шеннона.

double calculate\_entropy(const std::vector<int>& frequencies) {

double total\_count = std::accumulate(frequencies.begin(), frequencies.end(), 0);

double entropy = 0;

for (int freq : frequencies) {

if (freq != 0) {

double probability = freq / total\_count;

entropy -= probability \* log2(probability);

}

}

return entropy;

}

Листинг 1.1 – Метод для вычисления энтропии алфавита

Были созданы два файла с содержимым на македонском и турецком языках. Эти файлы считывались программой и исходя из их содержимого производились вычисления.

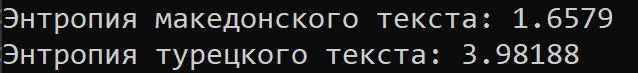


Рисунок 1.1 – Результаты вычисления энтропии македонского и турецкого алфавитов



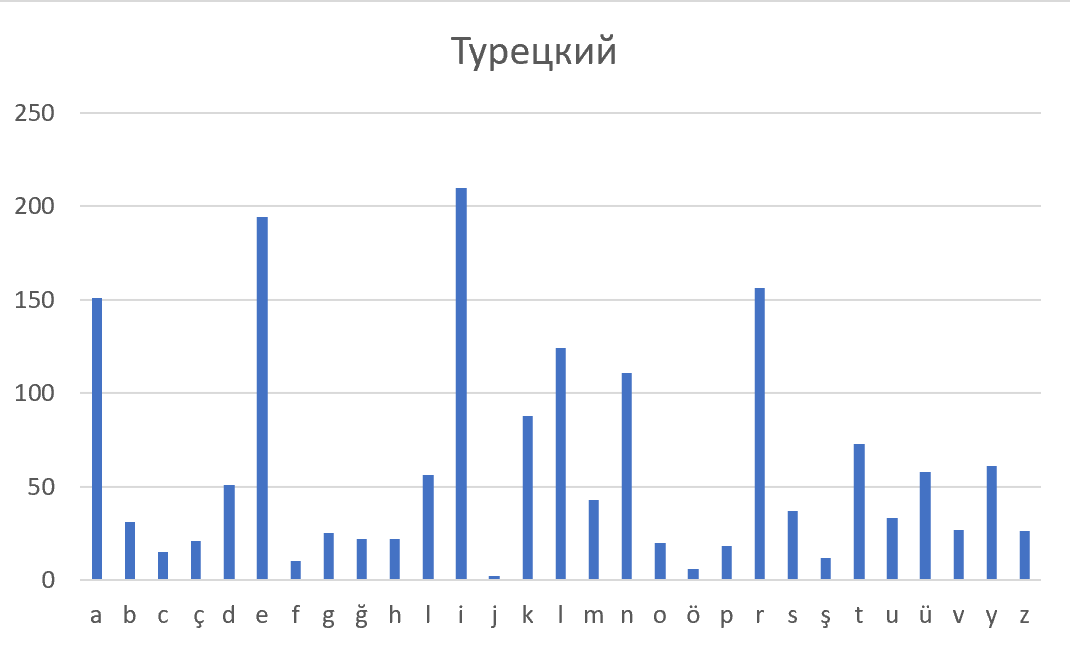


Рисунок 1.2 – Гистограммы частот появления символов

**Задание 2:** для входных документов, представленных в бинарных кодах, определить энтропию бинарного алфавита.

Для выполнения этого задания были созданы два файла для данных с 2 файлов в бинарном виде. Далее в программе эти файлы считывались и, используя тот же метод, высчитывалась энтропия.

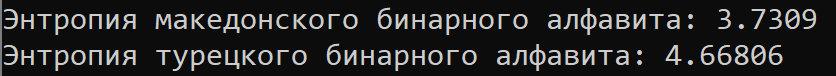


Рисунок 2.1 – Результаты вычисления энтропии бинарного алфавита

**Задание 3:** используя значения энтропии алфавитов, полученных в пунктах (а) и (б), подсчитать количество информации в сообщении, состоящем из собственных фамилии, имени и отчества (на основе исходного алфавита – (а) и в кодах ASCII – (б)); объяснить полученный результат.

Для вычисления количества информации в сообщении, состоящем из ФИО, каждая из энтропий была умножена на длину ФИО.

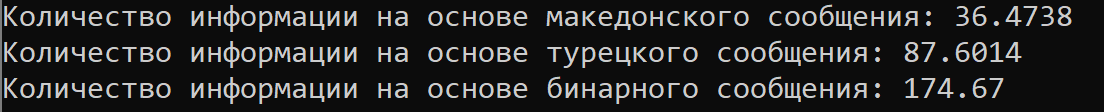


Рисунок 3.1 – Результаты вычисления количества информации для каждой из энтропий

В нашем случае, полученные значения количества информации отражают количества информации, необходимое для передачи этой информации с использованием указанных алфавитов.

**Задание 4:** выполнить задание пункта (в) при условии, что вероятность ошибочной передачи единичного бита сообщения составляет: 0,1; 0,5; 1,0.

Для выполнения этого задания были созданы два метода для вычисления эффективной энтропии и количества информации с вероятностью ошибочной передачи единичного бита.

double effective\_entropy(double p, double q, bool binary) {

if (binary && (p == 0 || q == 0)) {

return 1;

}

if (!binary) {

return 0;

}

return 1 - ((-p \* log2(p)) - (q \* log2(q)));

}

double amount\_info\_with\_error(double entropy, double p, double q, int count, bool binary) {

return entropy \* effective\_entropy(p, q, binary) \* count;

}

Листинг 4.1 – Методы для вычисления эффективной энтропии и количества информации с вероятностью ошибочной передачи единичного бита

В этом метода вычисляется эффективная энтропия алфавита и далее умножается на длину ФИО и энтропию алфавита для получения количества информации.

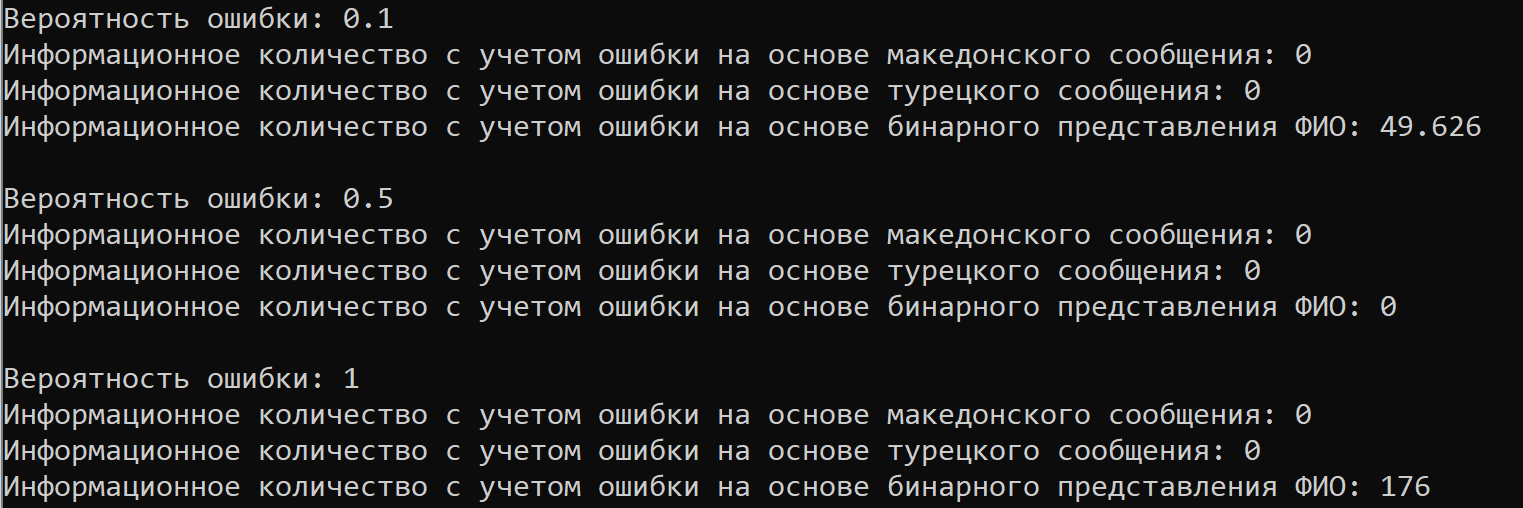


Рисунок 4.1 - Результат вычисления количества информации с вероятностью ошибочной передачи единичного бита

**Вывод:** Таким образом, чем выше вероятность ошибки, тем больше информации потребуется для надежной передачи сообщения, и в случае *P* *= 1.0*, передача становится невозможной, так как биты инвертируются. В небинарном алфавите при *P = 1* определить количество информации невозможно из-за того, что мощность алфавита больше 2.