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.

**Теоретические сведения**

Передача информации в ИС: **ИС → КП → ПС**.

**Информационный параметр сигнала** – отображение сообщения обеспечивается изменением какой-либо физической величины, характеризующей процесс (например, амплитуда, частота, фаза).

*Непрерывный (аналоговый) и дискретный сигналы.*

Информационный параметр **непрерывного сигнала** с течением времени может принимать любые мгновенные значения в определенных пределах.

**Дискретный** **сигнал** характеризуется конечным числом значений информационного параметра.

**Дискретные сообщения** состоят из последовательности дискретных знаков. (1,0) – двоичный или бинарный.

Построение сигнала по определенным правилам, обеспечивающим соответствие между сообщением и сигналом, называют **кодированием**. Преобразование сообщения в сигнал. Представление исходных знаков (символов) в другом алфавите с меньшим числом знаков.

**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.**

Nmin = 2, А = {0, 1} – минимум.

Один дискретный знак – **элементарное сообщение**, последовательность знаков – **сообщение**.

Если обозначить вероятность выбора каждого элемента алфавита p(аi), то

Двоичный канал – на основе двоичного алфавита. Если вероятность искажения переданного 0 = вероятности искажения переданной 1, то это двоичный симметричный канал.

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

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

Если все вероятности одинаковы и постоянны для всех символов алфавита, то:

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

Если вероятность ошибки в ДСК отлична от 0 (р > 0), переданное сообщение может содержать ошибки: Хk ≠ Yk. Количество информации в таком сообщении при его передаче по ДСК будет определяться не энтропией двоичного алфавита (в соответствии с выражением (2.3)), а эффективной энтропией Hе(A) алфавита или пропускной способностью канала:

где H(Y | X) – условная энтропия:

**Практическое задание**

а) 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);

Сперва необходимо задать алфавиты и рассчитать количество символов в сообщении, для этого используется функция calculateFrequencies (Рисунок 1.1).



Рисунок 1.1 – Функция для подсчета количества символов в сообщении

Далее считаем вероятность появления каждого символа в сообщении. Вероятность букв, которых в сообщении нет, не выводится. Для подсчета вероятности используется функция calculateProbabilities (Рисунок 1.2).

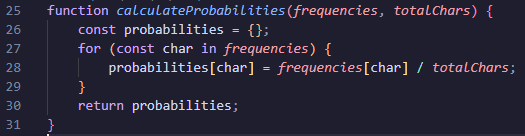


Рисунок 1.2 – Функция для подсчета вероятности

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

в) 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 энтропии используется функция calculateEntropy, код которой представлен на рисунке 1.3.

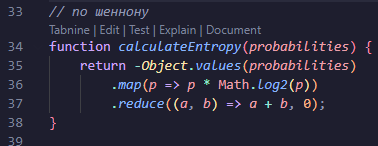


Рисунок 1.3 – Функция для вычисления энтропии

Для подсчета количества информации в сообщении используется функция calculateInformation.

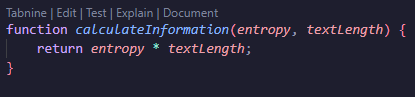


Рисунок 1.4 – Функция для подсчета количества информации

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

Для подсчета количества информации в сообщении с ошибкой используется функция calculateInformationWithError. Результаты работы на рисунках ниже.

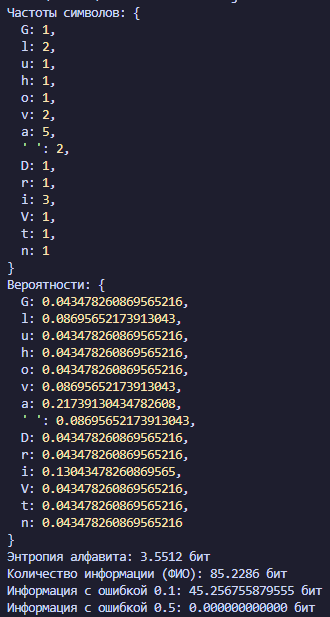
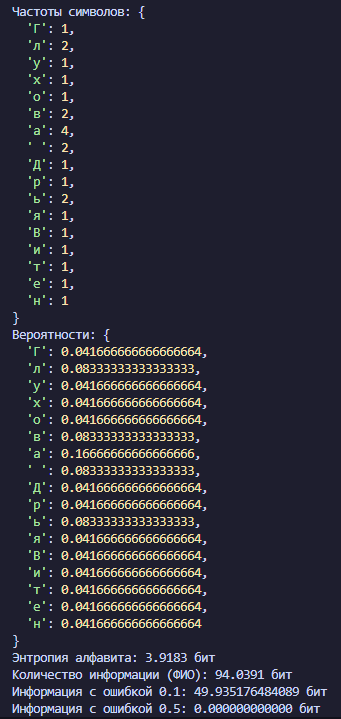


Рисунок 1.5 – Результат работы для русского и английского

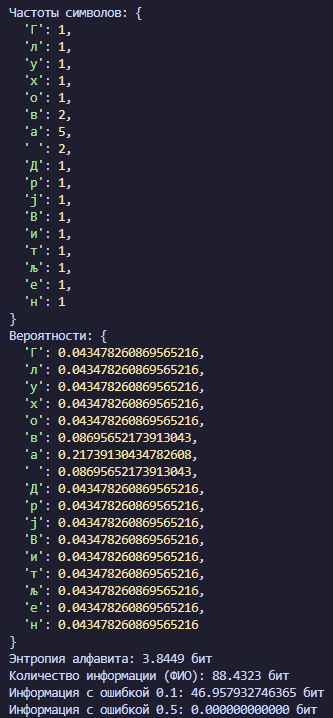


Рисунок 1.6 – Результат работы для сербского

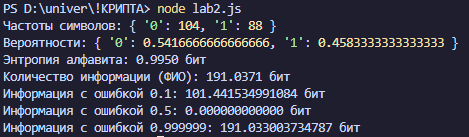


Рисунок 1.7 – Результат работы для бинарного

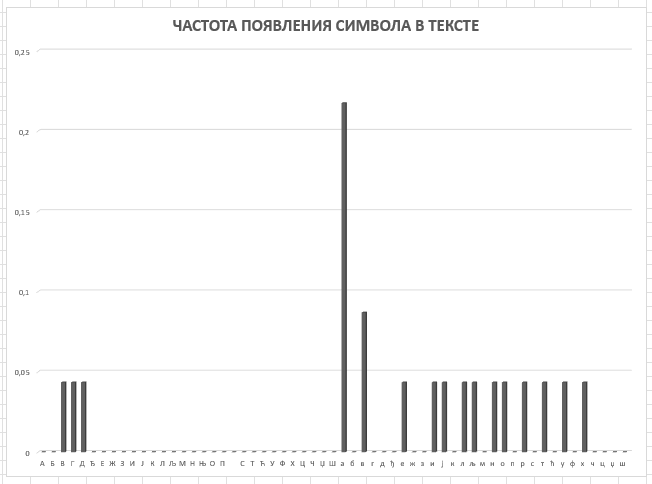


Рисунок 1.8 – График для сербского



Рисунок 1.9 – График для английского

**Вывод**

Данная лабораторная работа способствовала развитию практических навыков расчета и анализа параметров и информативных характеристик дискретных информационных систем, а также закреплению теоретических знаний по теории информации. В качестве практического результата разработано приложение для выполнения соответствующих расчетов.