Учреждение образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ  
ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информационных технологий

[Кафедра информационных](https://www.belstu.by/fakultety/fit/vm) систем и технологий

Специальность 1-40 01 01 Программное обеспечение информационных технологий

**Отчёт по лабораторной работе №2**

по дисциплине Информационная безопасность

Тема: Элементы теории информации. Информативность данных в различных кодировках

Исполнитель:

Студент 3 курса группы 6

Руководитель:

Ассистент Нистюк О. А.

Минск, 2024

**Лабораторная работа №2**

**Цель:** приобретение практических навыков трансформации данных и сопоставление энтропийных свойств используемых при этом алфавитов.

**Задачи:**

1. Закрепить теоретические знания по взаимной конвертации данных, представленных в кодах ASCII и base64.

2. Разработать приложение для конвертации произвольного документа в формат base64 и обратно.

3. Исследовать энтропийные характеристики используемых в конвертерах алфавитов.

4. Изучить особенности практической реализации операции XOR над данными, представленными в разных форматах.

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

# Теоретические сведения

Избыточностью алфавита называют уменьшение информационной нагрузки на один символ вследствие разной вероятности и взаимозависимости появления его символов в сообщениях.

В наиболее общем виде избыточность алфавита R можно оценить отношением энтропии по Хартли и по Шеннону:

R = [(HCh(A) – Hs(A)) / HCh(A)] \* 100%

Алгоритм base64 используется для кодирования бинарных данных.

Кодирование base64 разработано для представления произвольных последовательностей октетов в форме, позволяющей использовать строчные и прописные буквы. Используется 65-символьное подмножество набора символов US-ASCII, обеспечивающее представление одним печатным символом 6 битов данных (дополнительный 65-й символ используется для обозначения функции специальной обработки).

Процесс кодирования представляет группу из 24 последовательных битов в форме строки из 4 символов. Обработка выполняется слева направо, а 24-битная исходная группа образуется конкатенацией трех 8-битных групп (байтов). Данные 24 бита после этого трактуются как 4 сцепленных группы по 6 битов, каждая из которых транслируется в один символ алфавита base64.

Каждая 6-битная группа используется в качестве индекса массива из 64 печатных символов. Символы алфавита, соответствующие индексу, помещаются в выходную строку. Кодирование base64 с безопасным алфавитом используется для представления URL и имен файлов.

# Ход работы

Для конвертации текста документа в формат base64 была разработана функция base64Encode.

Функция base64Encode принимает текст документ и конвертирует его в формат base64. Код функции представлен на рисунке 2.1.

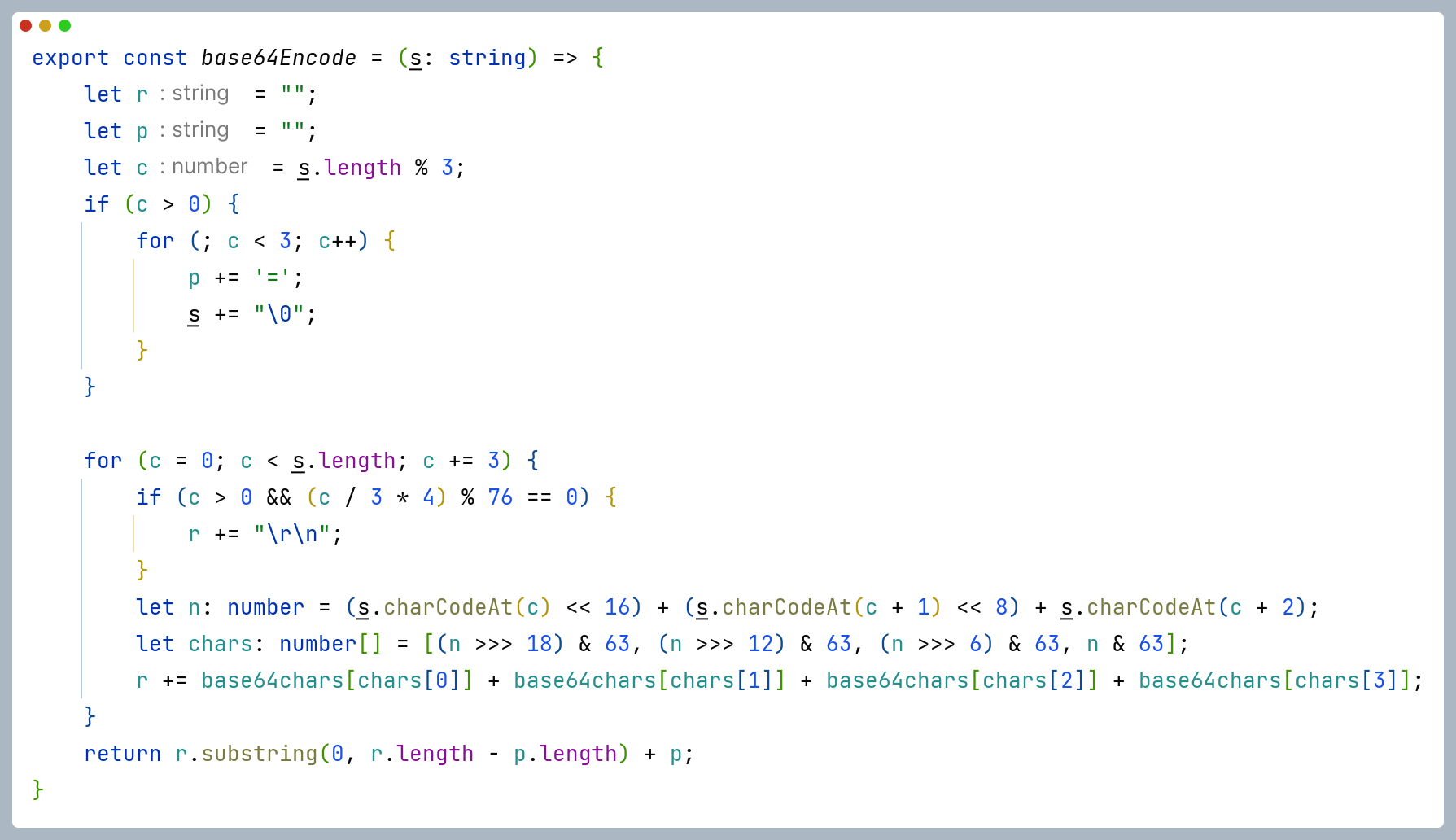


Рисунок 2.1 – Функция для конвертации текста в base64

В качестве входных данных был использован текстовый документ из предыдущей лабораторной работы на литовском языке. Результат конвертации представлен на рисунке 2.2.

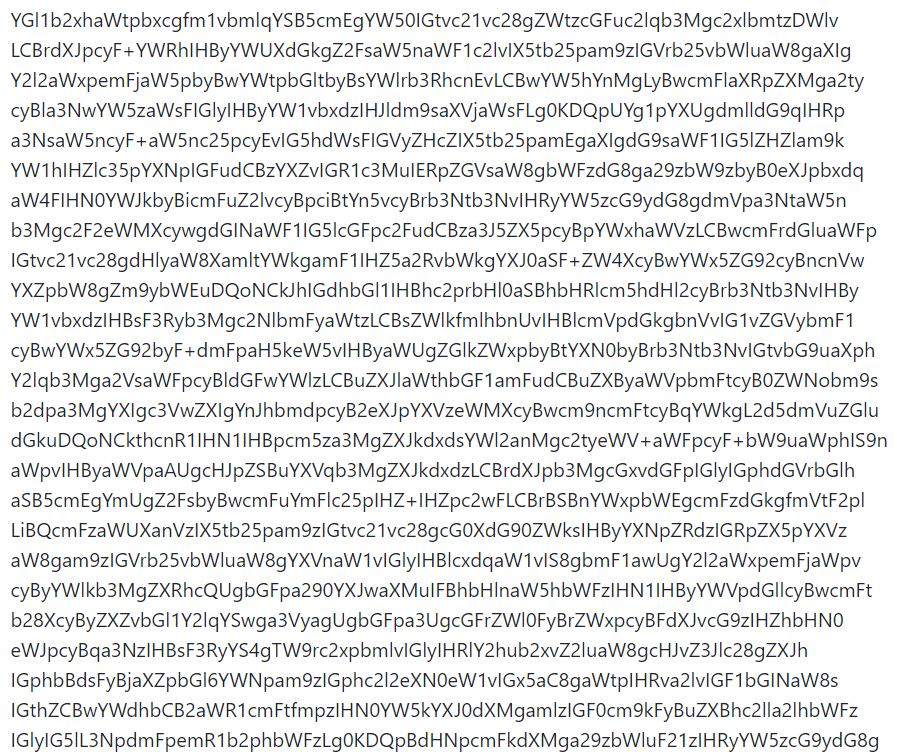


Рисунок 2.2 – Текст в формате base64

Далее, было необходимо получить частотные свойства алфавитов по документам. Для этого, с помощью приложения, созданного в предыдущей лабораторной работе, были построены гистограммы частот распределения символов в алфавитах. Гистограмма частот появления символов литовского алфавита представлена на рисунке 2.3.

Рисунок 2.3 – Гистограмма частот для литовского алфавита

Гистограмма частот появления символов алфавита base64 представлена на рисунке 2.4.

Рисунок 2.4 – Гистограмма частот для алфавита base64

На основе данных характеристик была определена энтропия Шеннона (определена с помощью приложения из предыдущей лабораторной работы) и энтропия Хартли. Для определения энтропии Хартли была создана функция calculateHartleyEntropy. Код функции представлен на рисунке 2.5.

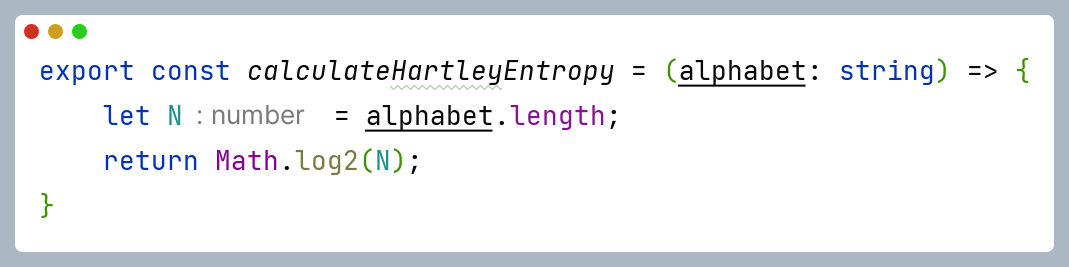


Рисунок 2.5 – Функция для определения энтропии Хартли

Кроме того, с помощью функции calculateRedundancy была определена избыточность алфавитов. Функция принимает значения энтропии Шеннона и энтропии Хартли. Код функции представлен на рисунке 2.6.

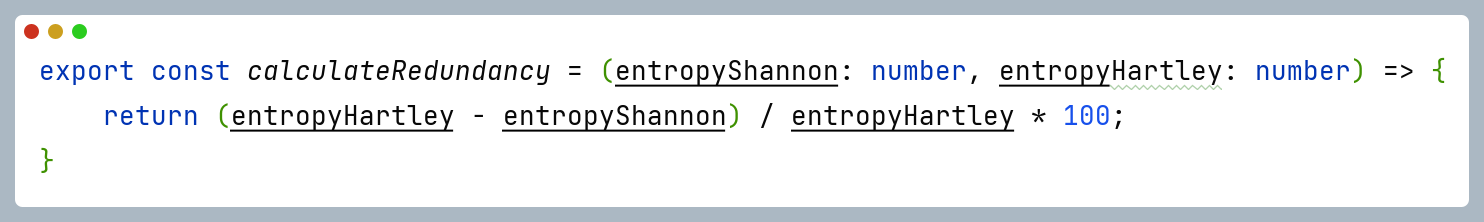


Рисунок 2.6 – Функция для определения избыточности алфавитов

Результат работы приложения представлен на рисунке 2.7.

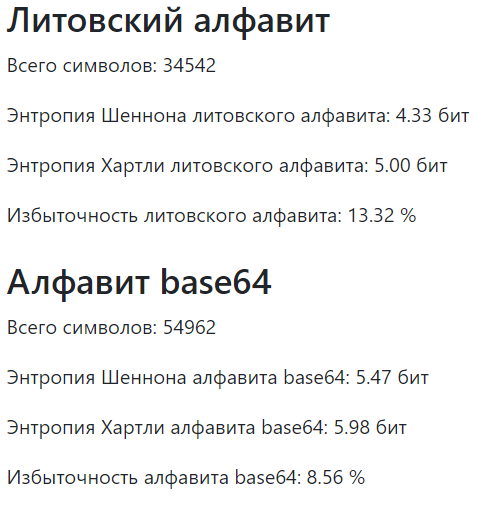


Рисунок 2.7 – Результат работы приложения

Далее нужно было разработать функцию, которая принимает в качестве аргументов два буфера (а и b) одинакового размера и возвращает XOR. Код функции представлен на рисунке 2.8.

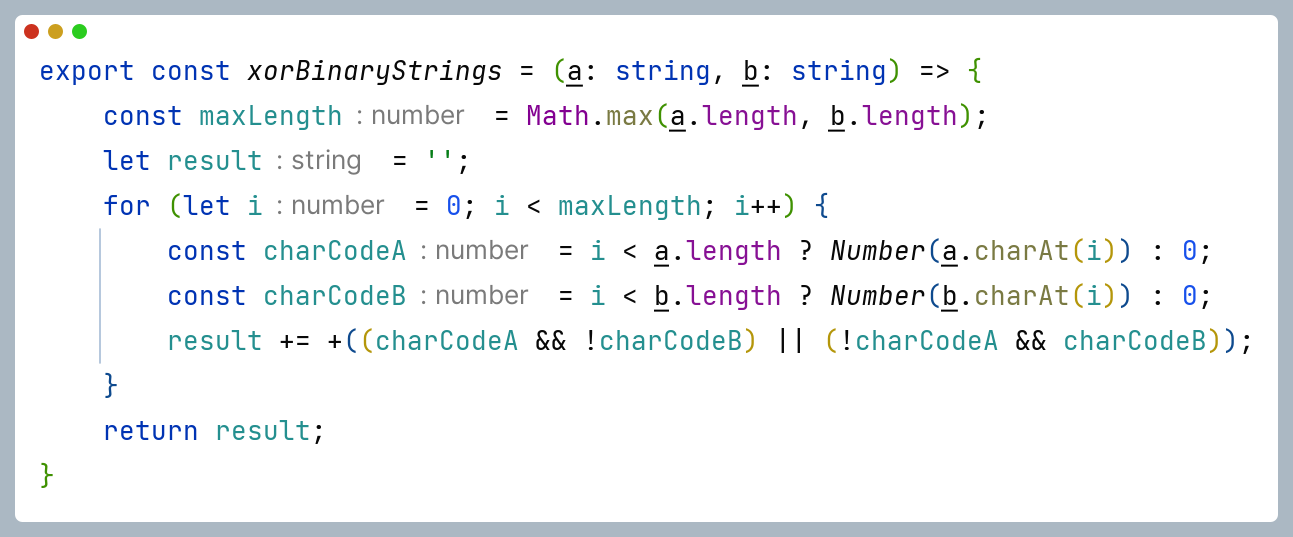


Рисунок 2.8 – Функция XOR для двух буферов

Результат работы функции для собственного имени и фамилии в кодах ASCII и base64 представлен на рисунке 2.9.

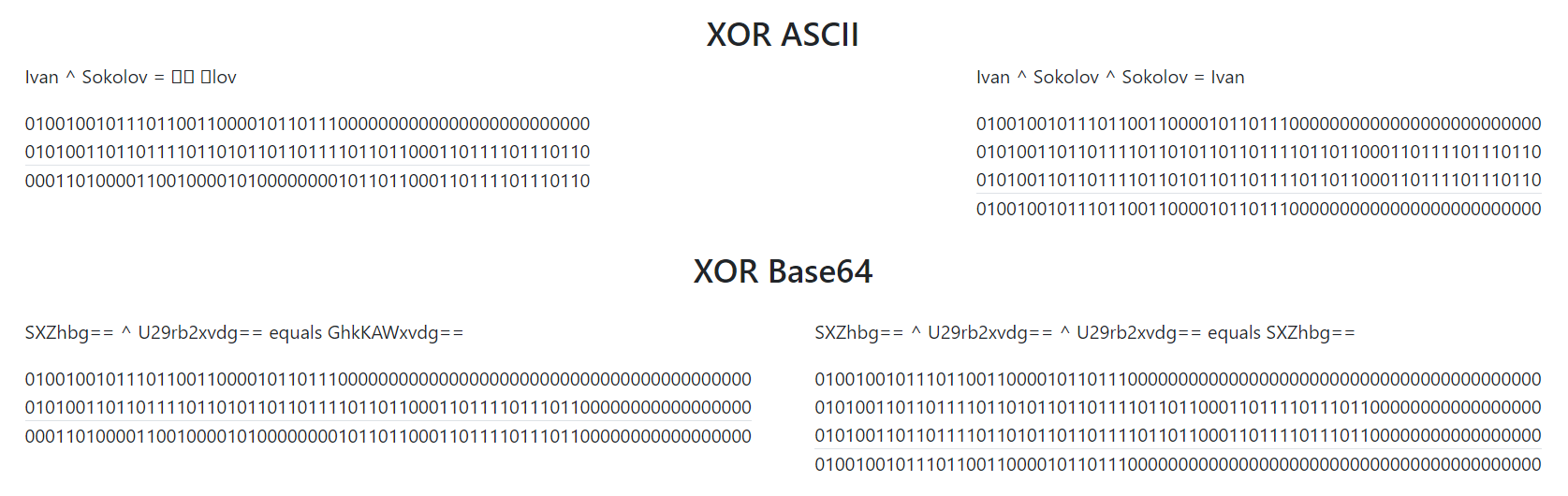


Рисунок 2.9 – Результат работы функции XOR

Можно заметить, что результатом операции *а*XOR*b*XOR*b* будет а.

**Вывод:** в ходе выполнения лабораторной работы было разработано приложение для конвертации текста на латинице в документ формата base64. Кроме того, была определена энтропия Шеннона, энтропия Хартли и избыточность литовского алфавита и алфавита base64. Также были изучены особенности реализации операции XOR над данными, представленными в разных форматах (ASCII и base64).