=НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.04 «Программная инженерия»

Дисциплина «Информатика»

**Отчет**

По лабораторной работе №4

*Исследование протоколов,*

*форматов обмена информацией и языков разметки*

*документов*

Вариант 35

Студент

Грищенко Андрей Викторович

*P3108*

Преподаватель

Рудникова Тамара Владимировна

Санкт-Петербург 2023 г.

Оглавление

[Задание 3](#_heading=h.gjdgxs)

[Основные этапы вычисления 3](#_heading=h.30j0zll)

[Обязательное задание 3](#_heading=h.1fob9te)

[Дополнительное задание №1 3](#_heading=h.3znysh7)

[Дополнительное задание №2 3](#_heading=h.2et92p0)

[Дополнительное задание №3 3](#_heading=h.tyjcwt)

[Дополнительное задание №4 3](#_heading=h.3dy6vkm)

[Дополнительное задание №5 3](#_heading=h.1t3h5sf)

[Заключение 3](#_heading=h.4d34og8)

[Источники 3](#_heading=h.2s8eyo1)

# Задание

Обязательное задание: написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый путём простой замены метасимволов исходного формата на метасимволы результирующего формата.

Дополнительное задание №1

a) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.

b) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.

c) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

Дополнительное задание №2

a) Переписать исходный код, добавив в него использование

регулярных выражений.

b) Сравнить полученные результаты и объяснить их

сходство/различие. Объяснение должно быть отражено в

отчёте.

Дополнительное задание № 3

а) Переписать исходный код таким образом, чтобы для решения задачи использовались формальные грамматики. То есть ваш код должен уметь осуществлять парсинг и конвертацию любых данных, представленных в исходном формате, в данные, представленные в результирующем формате: как с готовыми библиотеками из дополнительного задания №1.

b) Проверку осуществить как минимум для расписания с двумя учебными днями по два занятия в каждом.

с) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

Дополнительное задание № 4

a) Используя свою исходную программу из обязательного задания и программы из дополнительных заданий, сравнить стократное время выполнения парсинга + конвертации в цикле.

b) Проанализировать полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

Дополнительное задание № 5

a) Переписать исходную программу, чтобы она осуществляла парсинг и конвертацию исходного файла в любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.

b) Проанализировать полученные результаты, объяснить особенности использования формата. Объяснение должно быть отражено в отчёте.

# Основные этапы вычисления

## Обязательное задание



Рисунок 1. Программное решение обязательного задания на языке Python

## Дополнительное задание №1



Рисунок 2. Программное решение дополнительного задания №1 на языке Python

## Дополнительное задание №2



Рисунок 3. Программное решение дополнительного задания №2 на языке Python

## Дополнительное задание №3



Рисунок 4. Код основного файла для решения дополнительного задания №3 на Python

На рис. 4 представлен код файла main.py, в котором мы используем собственную библиотеку.



Рисунок 5. Код класса узлов, реализующий конвертацию файлов

На рис. 5 представлен код файла node.py с классом Node, позволяющий выполнить парсинг XML файла в “узлы”, строя граф в памяти, а также сохранение его в формате YAML



Рисунок 6. Код класса-интерфейса для парсинга файлов

На рис. 6 представлен класс Parser, выполняющий функцию интерфейса к классу Node, поскольку реализует некоторые тонкости класса узлов, а также не имеет вспомогательных функций.

## Дополнительное задание №4



Рисунок 7. Программное решение дополнительного задания №4 на языке Python

Результаты:

0: 0.00038213300013012486

1: 0.06325255299998389

2: 0.0004555300001811702

3: 0.0010766890000013518

Пояснение: 0 и 2 пункт не сильно различаются, поскольку первый только производит замену, а второй использует регулярные выражения. Довольно большое время в 1 пункте, поскольку для перевода используется две библиотеки: для парсинга XML файла в словарь, и конвертирование его в YAML. В 3 пункте сравнительно больше времени, так как код не оптимизирован

## Дополнительное задание №5

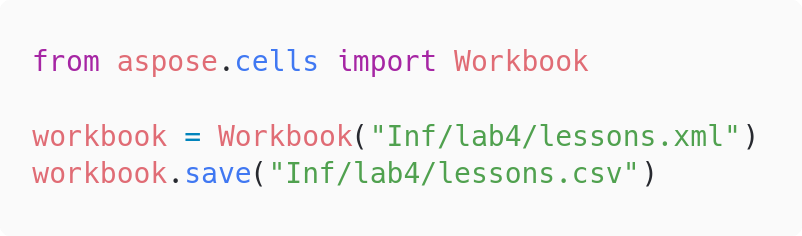


Рисунок 8. Программное решение дополнительного задания №5 на языке Python

# Заключение

В ходе проделанной лабораторной работы я изучил что такое формальные грамматики, углубился в языки разметки, а также написал несколько вариантов парсинга

# Источники

Хабр – URL: <https://habr.com/ru/articles/>

Википедия – URL: [https://ru.wikipedia.org/wiki/Заглавная\_страница](https://ru.wikipedia.org/wiki/%D0%97%D0%B0%D0%B3%D0%BB%D0%B0%D0%B2%D0%BD%D0%B0%D1%8F_%D1%81%D1%82%D1%80%D0%B0%D0%BD%D0%B8%D1%86%D0%B0)