Федеральное государственное автономное образовательное учреждение высшего

Образования «Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.04 Программная инженерия

Дисциплина «Информатика»

**Лабораторная работа №4**

**Исследование протоколов,**

**форматов обмена информацией и языков разметки документов**

**Вариант 1**

Работу выполнил:

Бондаренко Артем Андреевич

Группа P3110

Преподаватель, принимающий работу:

Малышева Татьяна Алексеевна

Доцент факультета ПИиКТ

г. Санкт-Петербург 2023 г.

**Содержание**

[Задание 3](#_Toc152756141)

[Основные этапы вычисления 4](#_Toc152756142)

[Обязательное задание 4](#_Toc152756143)

[**Дополнительное задание №1** 4](#_Toc152756144)

[**Дополнительное** **задание №2** 5](#_Toc152756145)

[**Дополнительное задание №4** 5](#_Toc152756146)

[Вывод 6](#_Toc152756147)

[Список используемой литературы 7](#_Toc152756148)

# 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

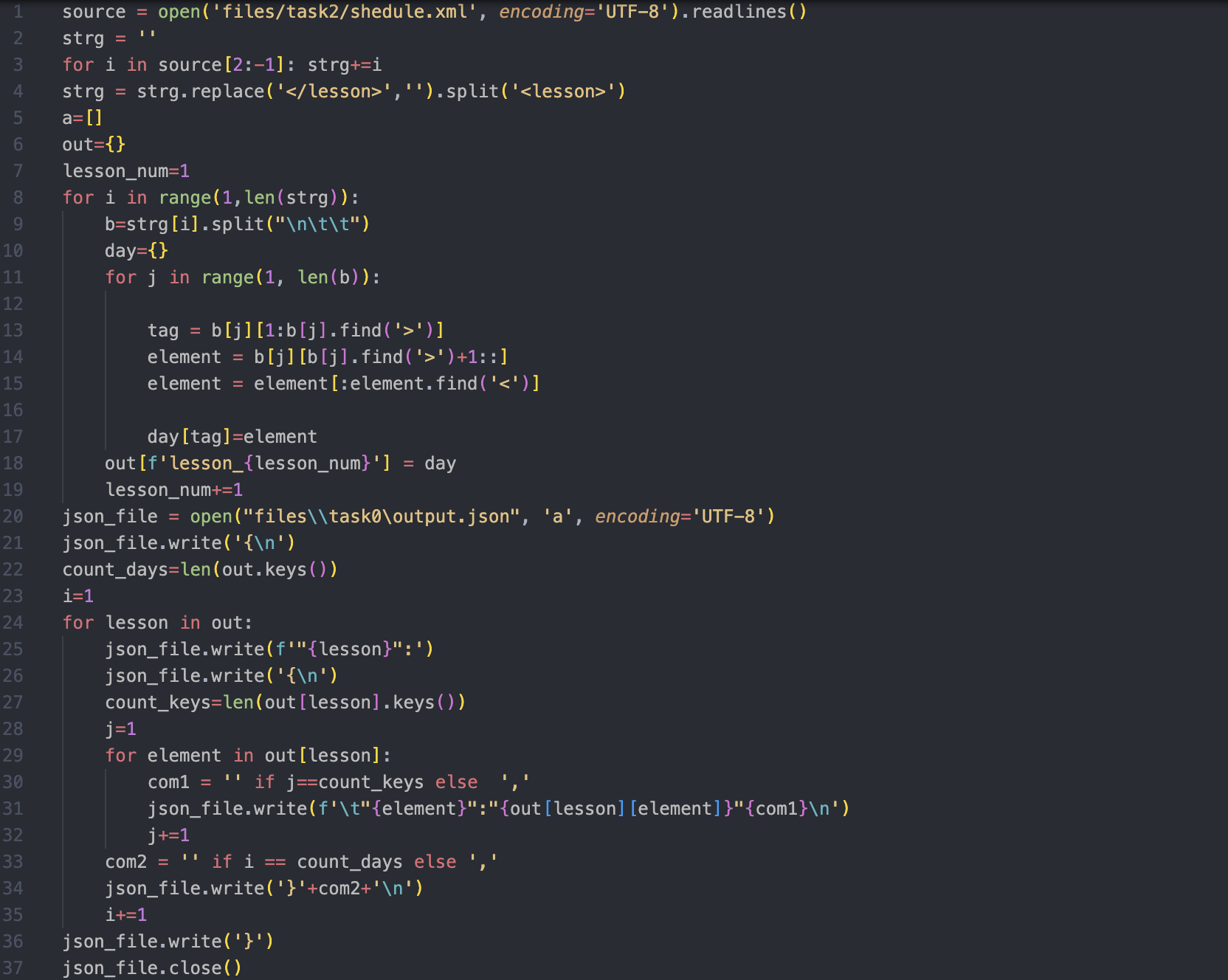
a) Переписать исходную программу, чтобы она осуществляла парсинг и конвертацию исходного файла в любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.

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

# Основные этапы вычисления

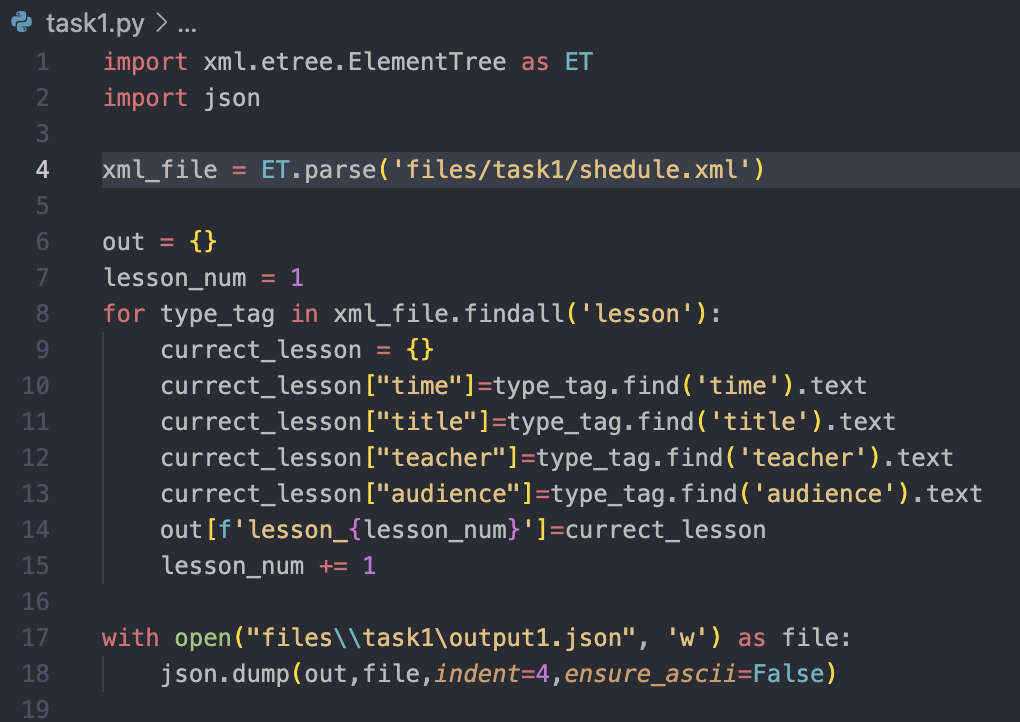
## Обязательное задание

Файл task0.py

****

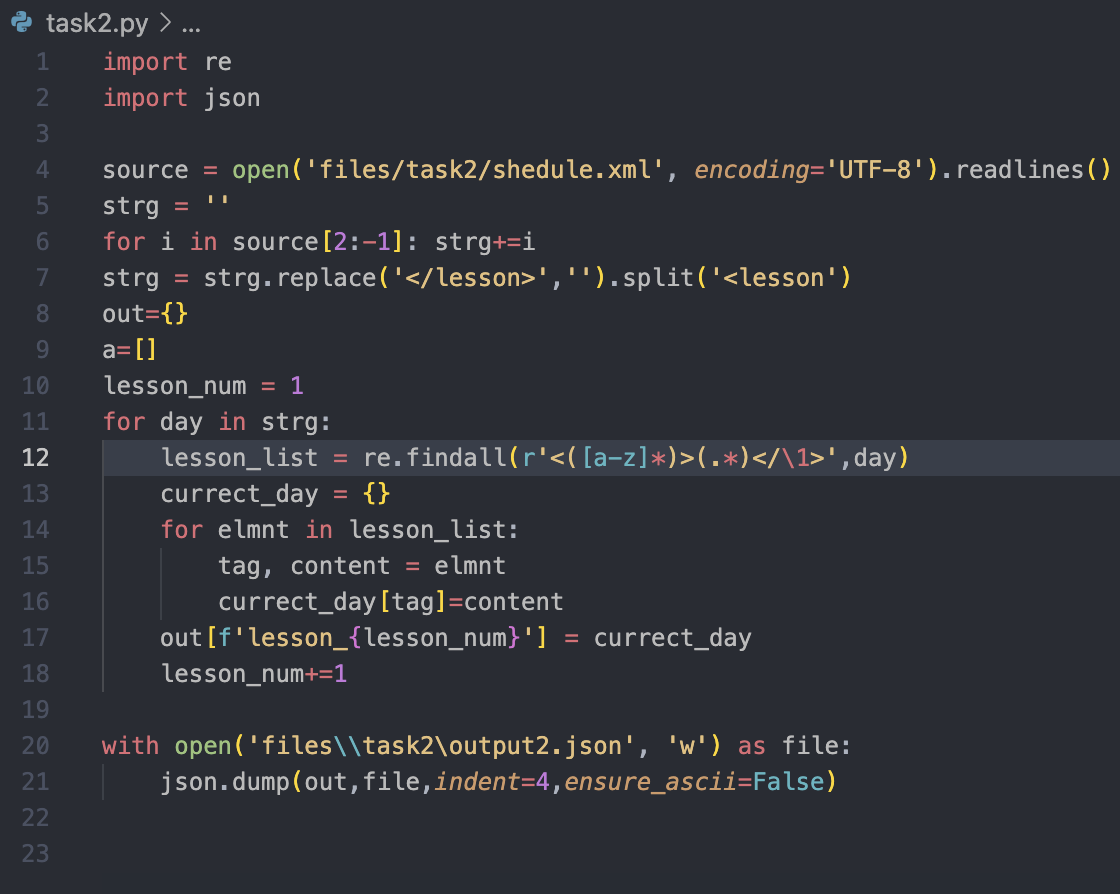
## **Дополнительное задание №1**

Файл task1.py

****

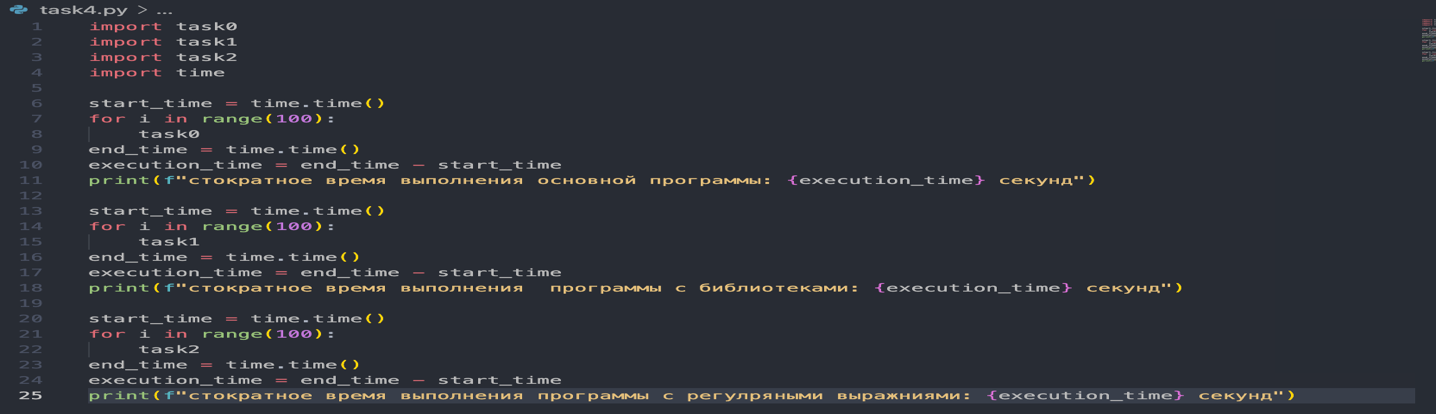
## **Дополнительное** **задание №2**

Файл task2.py



## **Дополнительное задание №4**

Файл task4.py



Результат выполнения работы файла task4.py:

стократное время выполнения основной программы: 6.152557373046875e-06 секунд

стократное время выполнения программы c библиотеками: 7.198883056640625e-06 секунд

стократное время выполнения программы с регулряными выражниями: 6.9141387939453125e-06 секунд

Пояснение:

Конвертация файлов своим решением проходит быстрее, чем конвертация при

Помощи использования больших библиотек, так как мы рассматриваем конкретный случай(при выполнении основного задания мы опираемся лишь на созданный нами файл), тем самым, нам не надо искать или рассматривать множество различных ситуаций и метасимволов. Конвертация при помощи использования регулярных выражений так же выполняется довольно долго , так как работа с ними ориентирована на большие объемы входных данных.

# Вывод

В ходе выполнения данной лабораторной работы я узнал, что такое парсинг, научился осуществлять данную процедуру, выявлять основные ошибки, которые могут возникнуть при данном процессе, и избегать их.

# Список используемой литературы

1. Что такое XML: <https://habr.com/ru/articles/524288/>
2. YAML: <https://ru.wikipedia.org/wiki/YAML>
3. Формальная грамматика: <https://ru.wikipedia.org/wiki/Формальная_грамматика>