Федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

**Отчет**

**по Лабораторной работе No4**

**«Исследование протоколов, форматов обмена информацией и языков разметки документов»**

**Вариант 8**

Выполнил: Киеу Чыонг Занг

Группа: P3107

Преподаватель: Белозубов Александр Владимирович

Г.Санкт-Петербург

2022

Table of Contents

[Порядок выполнения работы 2](#_Toc121667772)

[Выполнение 4](#_Toc121667773)

[1. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного 4](#_Toc121667774)

[2. написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новы 5](#_Toc121667775)

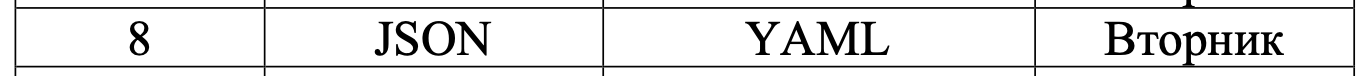
[3. Дополнительное задание задание №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную). 7](#_Toc121667776)

[4. Дополнительное задание задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную). 8](#_Toc121667777)

[5 .Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную). 11](#_Toc121667778)

[6. Дополнительное задание задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную. 11](#_Toc121667779)

[Заключение: 13](#_Toc121667780)



# Порядок выполнения работы

1. Определить номер варианта как остаток деления на 36 порядкового номера в списке группы в ISU. В случае, если в данный день недели нет занятий, то увеличить номер варианта на восемь.

2. Изучить форму Бэкуса-Наура.

3. Изучить особенности протоколов и форматов обмена информацией между системами: JSON, YAML, XML.

4. Понять устройство страницы с расписанием для своей группы: https://itmo.ru/ru/schedule/0/P3110/schedule.htm

5. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного.

6. Обязательное задание (позволяет набрать до 65 процентов от максимального числа баллов БаРС за данную лабораторную): написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.

7. Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов.

8. Дополнительное задание задание №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

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.

9. Дополнительное задание задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

a) Переписать исходный код, добавив в него использование регулярных выражений.

b) Сравнить полученные результаты и объяснить их сходство/различие.

10.Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

a) Используя свою исходную программу из обязательного задания, программу из дополнительного задания №1 и программу из дополнительного задания №2, сравнить десятикратное время выполнения парсинга + конвертации в цикле.

b) Проанализировать полученные результаты и объяснить их сходство/различие.

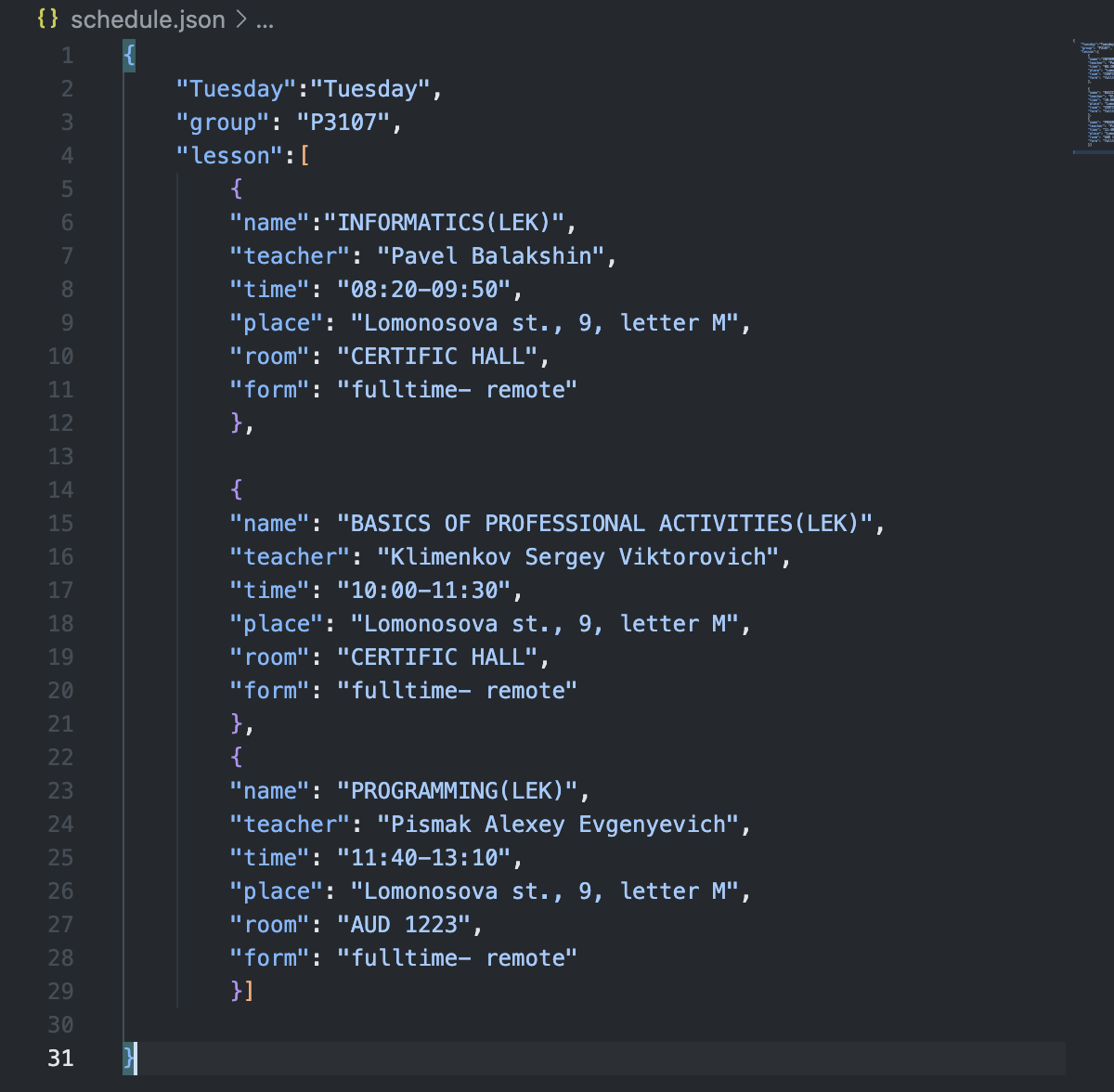
11.Дополнительное задание задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную.

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

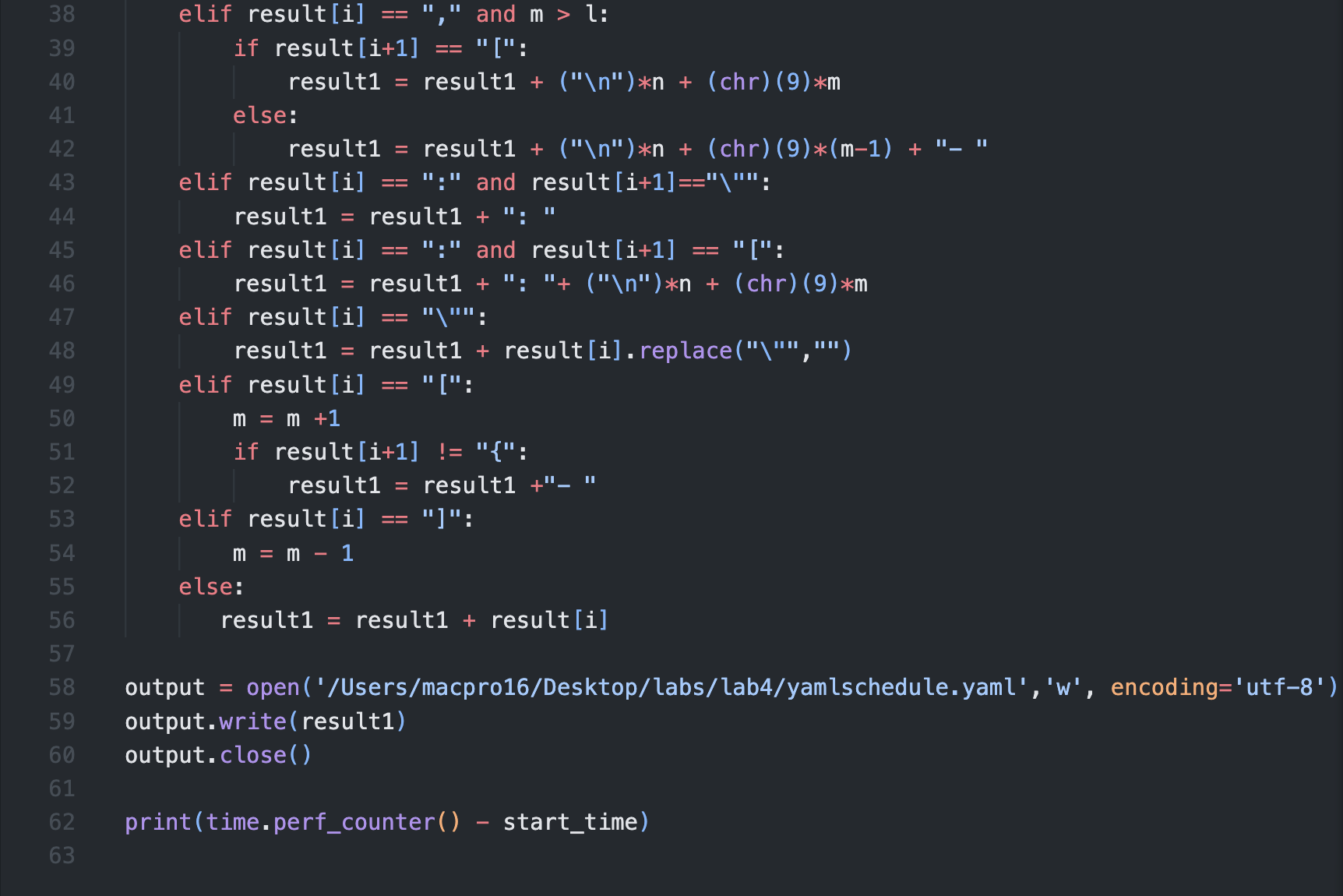
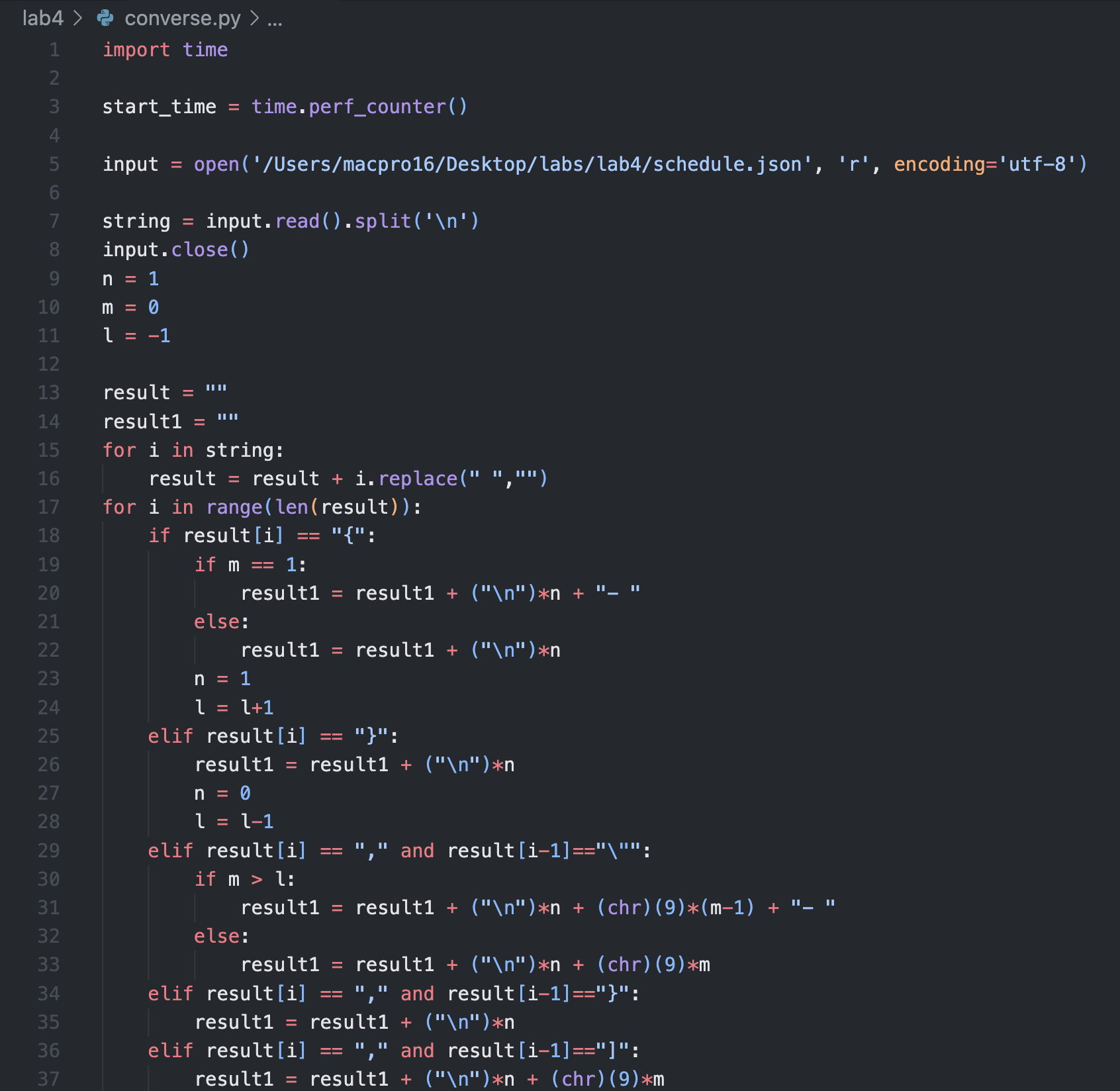
b) Проанализировать полученные результаты, объяснить осоебнности использованного формата.

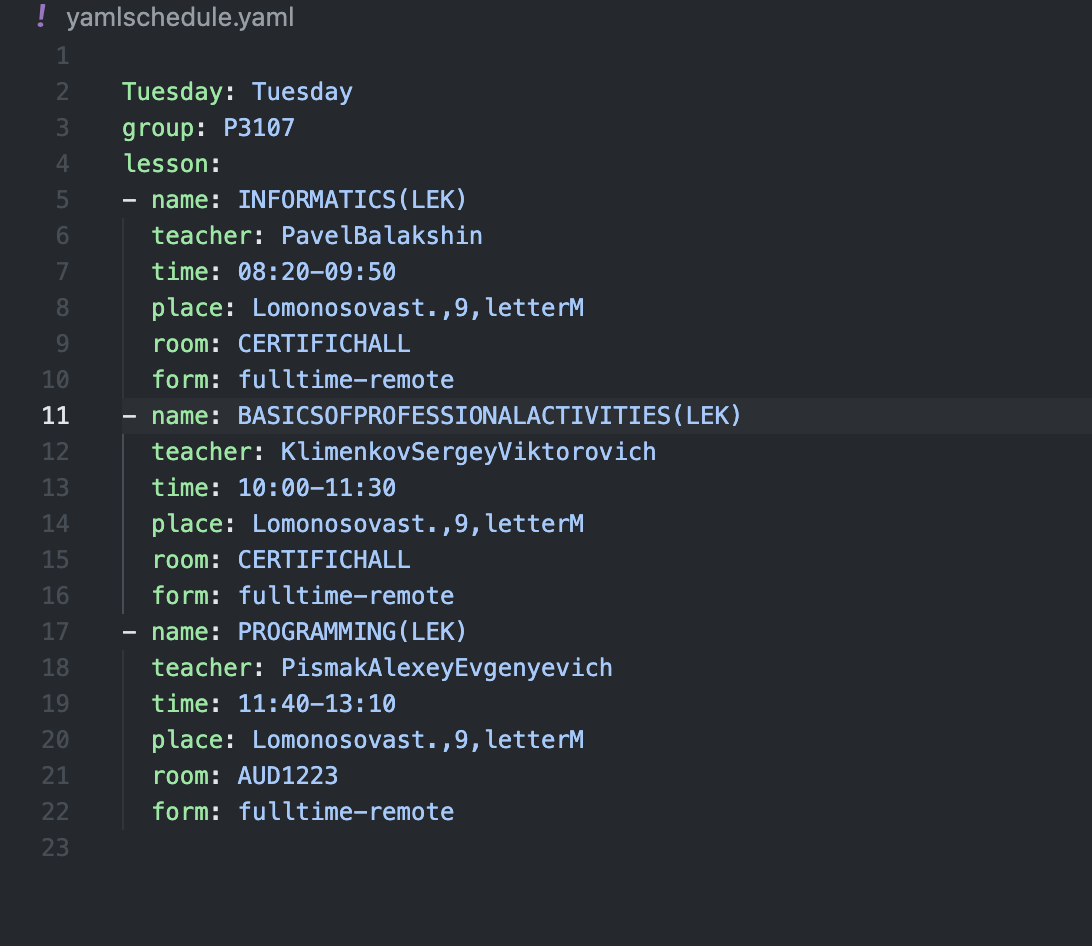
Выполнение

## 1. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного



2. написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новы



**yamlschedule.yaml**

## 3.Дополнительное задание задание №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

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.

Используемая библиотека: yaml, json.



**toyaml.yaml**



\* Сравните: почти такой же, как и предыдущий результат из-за небольшой разницы в алгоритме смены персонажа.

## 

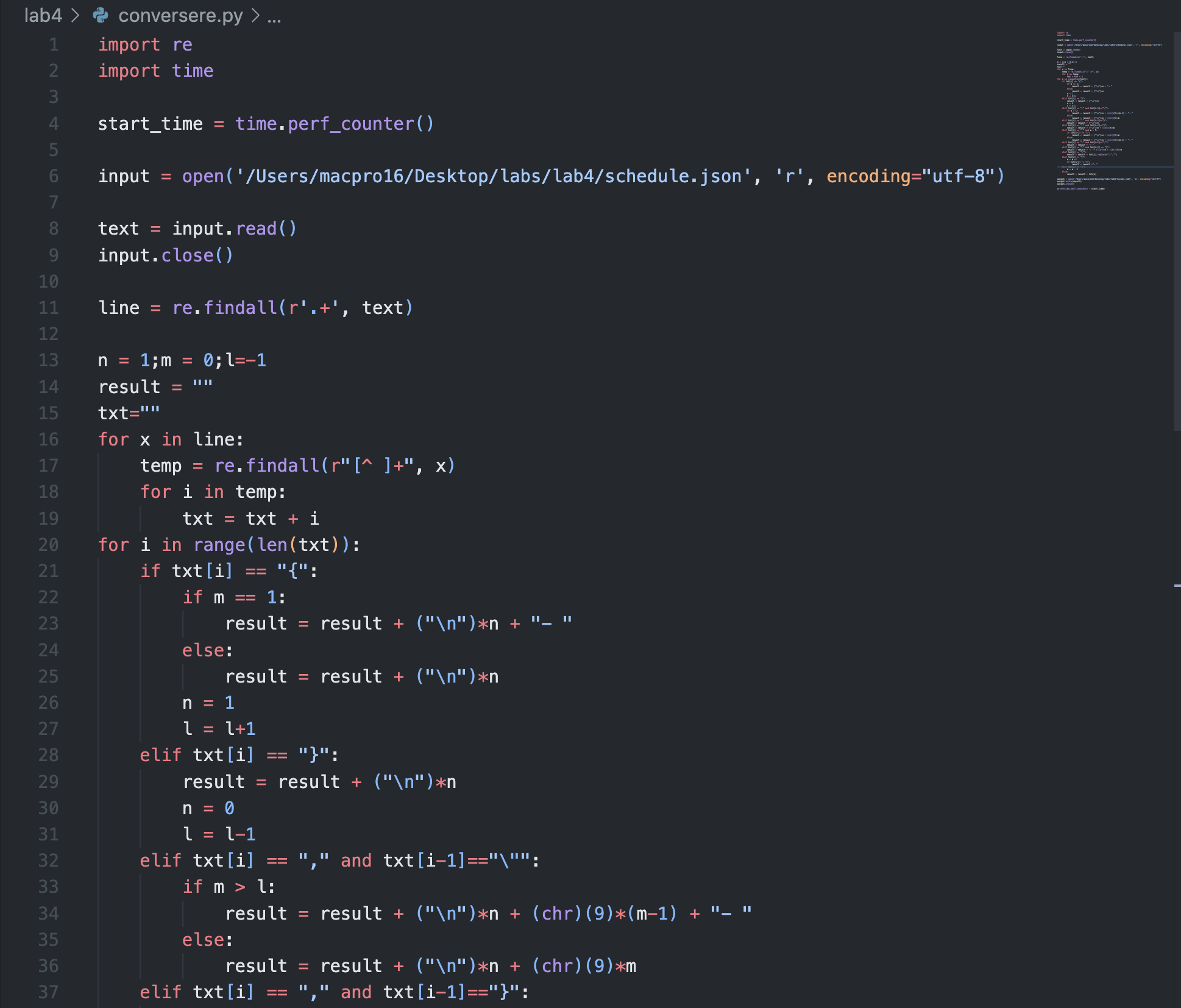
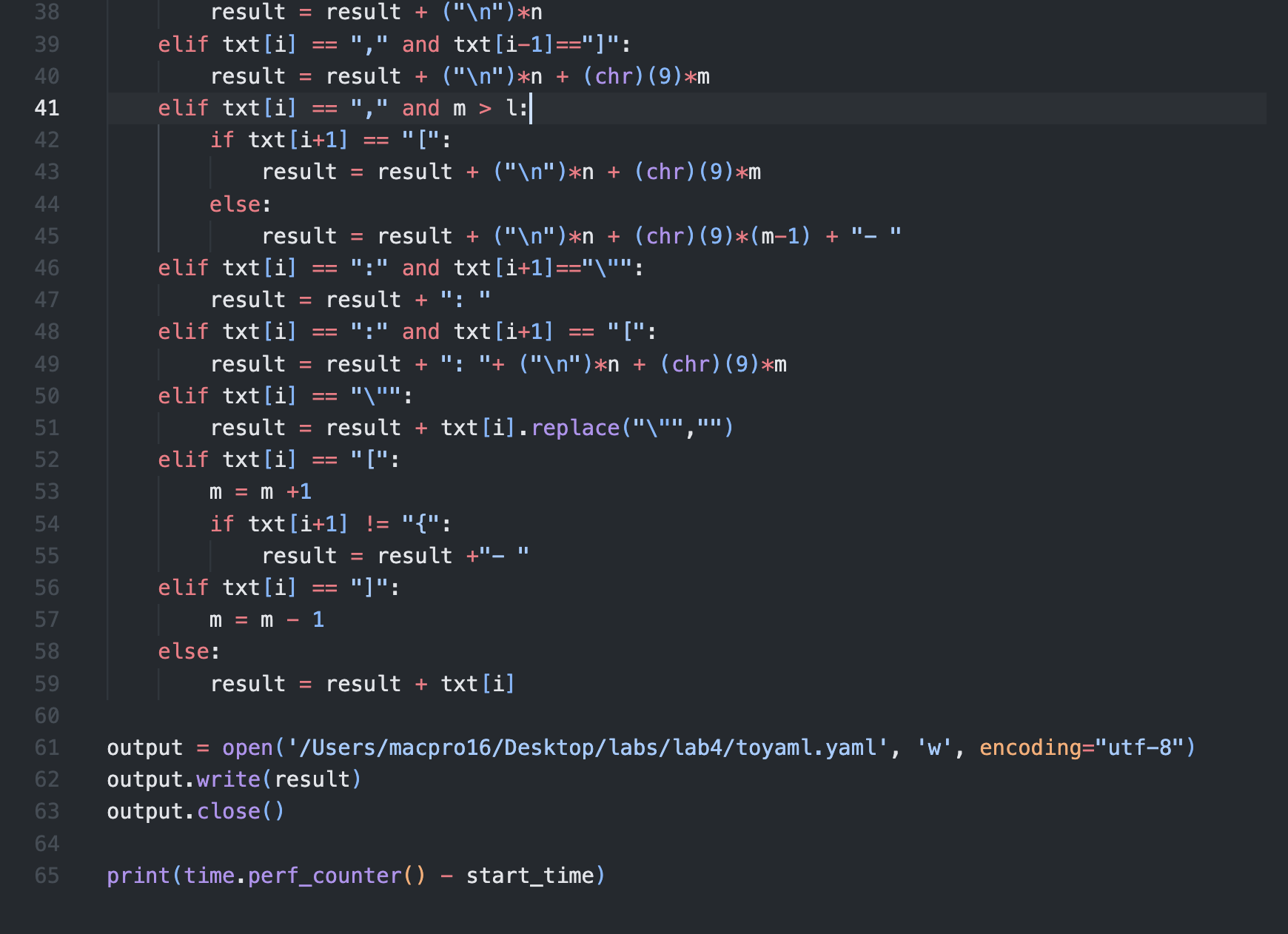
## 4. Дополнительное задание задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

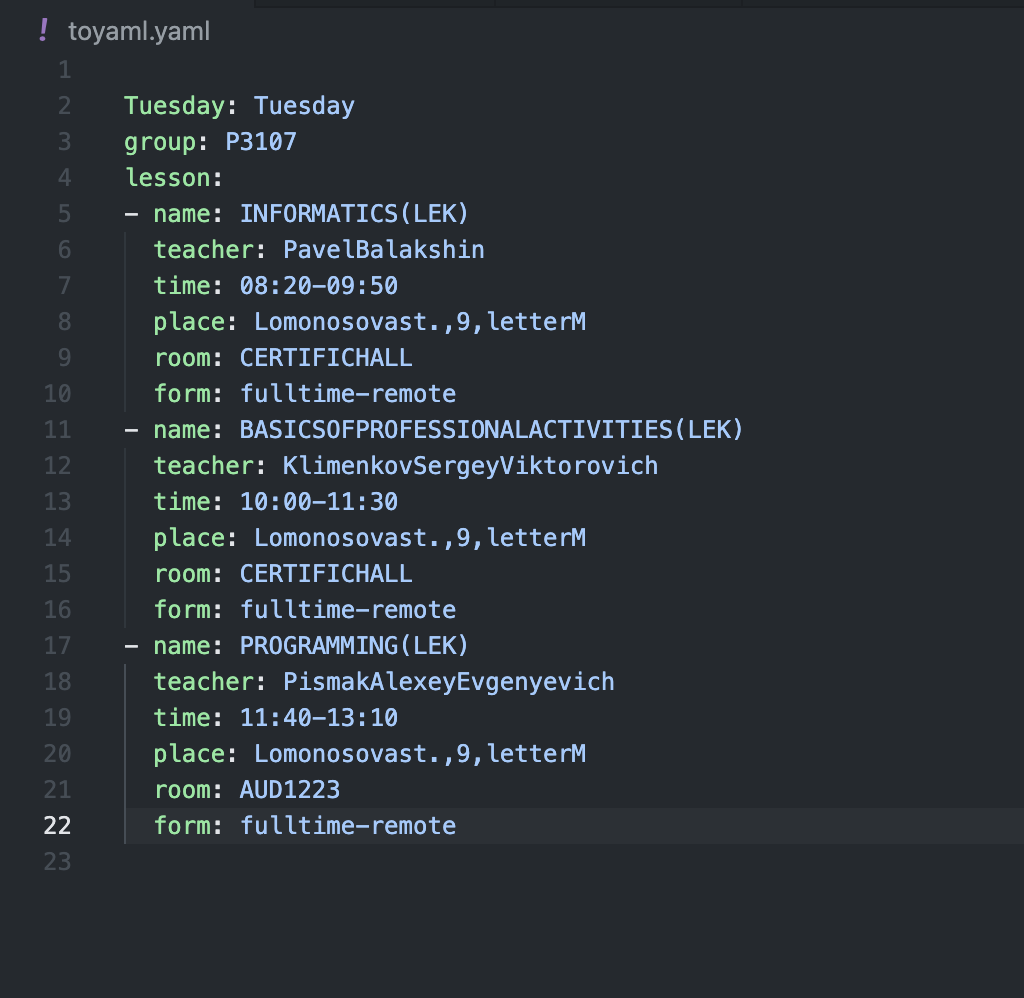
a) Переписать исходный код, добавив в него использование

регулярных выражений.

b) Сравнить полученные результаты и объяснить их

сходство/различие.



**toyaml.yaml**

\* Сравните: То же, что и первый результат, но немного отличается от второго результата из-за небольшой разницы в алгоритме смены персонажа.

## 5 .Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

a) Используя свою исходную программу из обязательного

задания, программу из дополнительного задания №1 и

программу из дополнительного задания №2, сравнить

десятикратное время выполнения парсинга + конвертации в цикле.

-среднее время выполнения 100 раз первой программы: 0.00029533375

-среднее время выполнения 100 раз второй программы: 0.00076650625

-среднее время выполнения 100 раз третье программы: 0.00037325625

b) Проанализировать полученные результаты и объяснить их

сходство/различие.

-Первая программа является самой быстрой, поскольку не использует никаких дополнительных внешних библиотек.

-Третья программа использует регулярное выражение, поэтому она немного медленнее, но во многих случаях регулярное выражение помогает нам писать код быстрее.

- Вторая программа использует встроенную библиотеку и написана для работы с многими различными случаями, в то время как мой код скорее всего не будет работать ни на чем, кроме моего одного конкретного кейса.

Поэтому обычно, если время выполнения не слишком важно, я бы предпочел использовать внешнюю библиотеку, потому что она написана для большинства случаев и очень проста в использовании.

## 6. Дополнительное задание задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную.

a) Переписать исходную, чтобы она осуществляла парсинг и

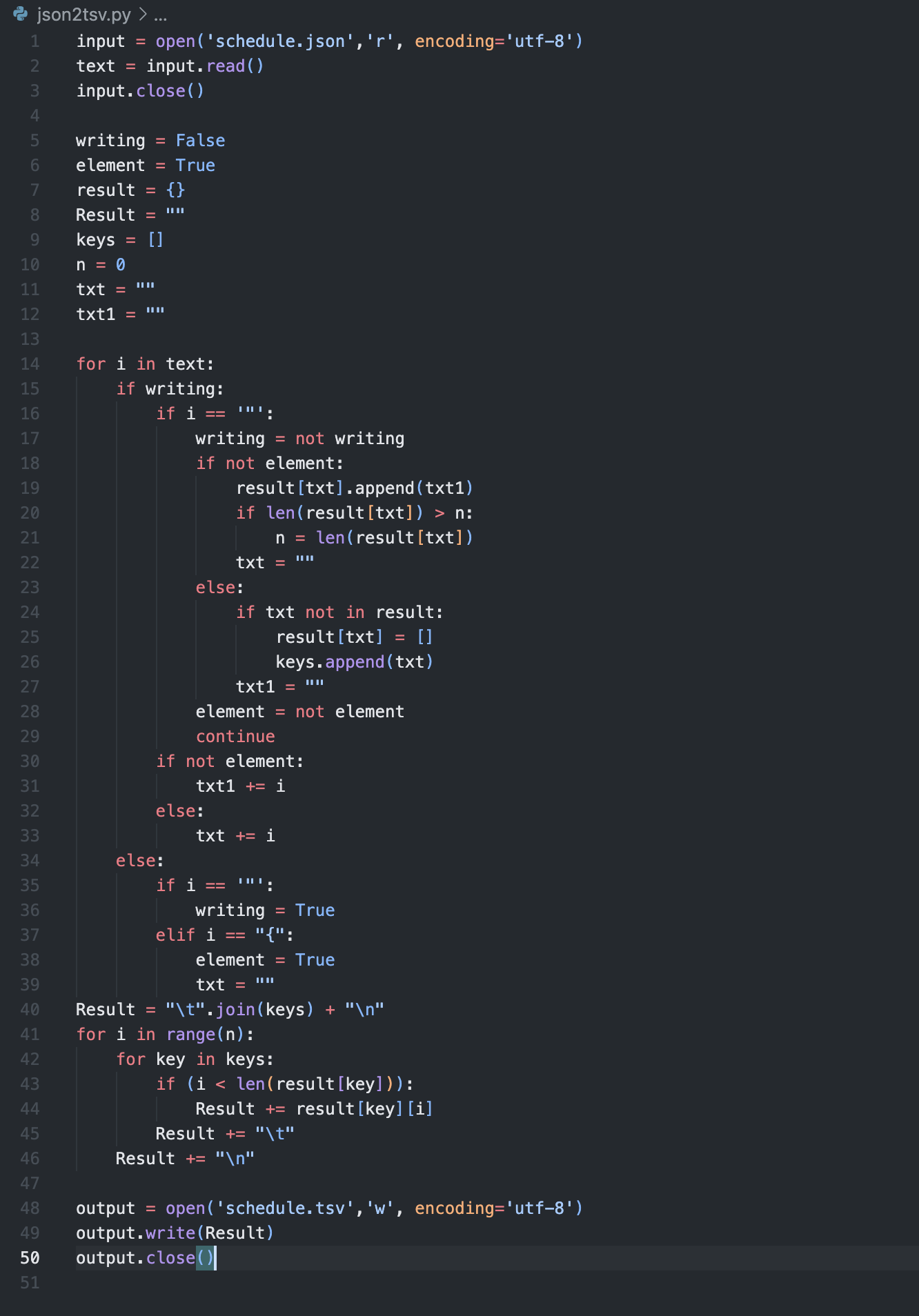
конвертацию исходного файла в любой другой формат (кроме

JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и

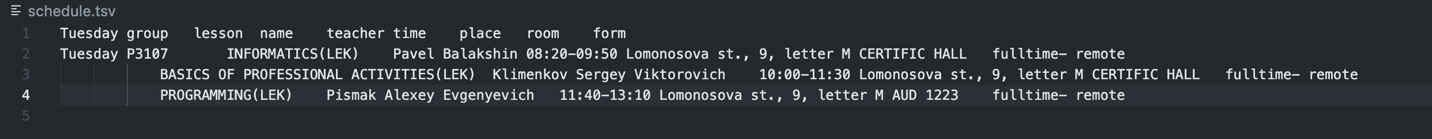
т.п.

b) Проанализировать полученные результаты, объяснить

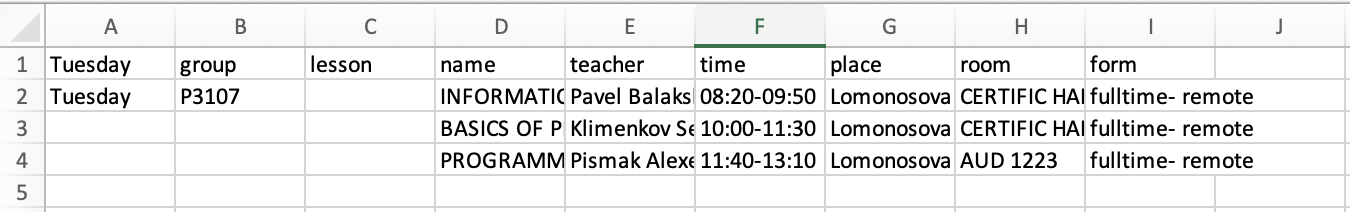
осоебнности использованного формата



**schedule.tsv**



**schedule.tsv открывается в Excel**

****

Особенности

TSV используется для таблиц и в нем тяжело задавать массивы и объекты. Я поместил все ключи в первую строку, а их значения в столбик под ними.

# Заключение

После выполнения этой лабораторной работы я изучил форматы данных на компьютере, как организовать данные JSON, XML, YAML и другие. Я уже знаю, как конвертировать данные между форматами.