Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»

Факультет Программной Инженерии и Компьютерной Техники

**Лабораторная работа по информатике №4**

Вариант: 15

Преподаватель: Рудникова Тамара Владимировна

Выполнил: Терехин Никита Денисович

Группа: Р3108

Санкт-Петербург

2022г

Оглавление

[Задание 1 3](#_Toc118890984)

[Задание 2 6](#_Toc118890985)

[Задание 3 8](#_Toc118890986)

[Задание 4 9](#_Toc118890987)

[Вывод 9](#_Toc118890988)

[Список литературы. 10](#_Toc118890989)

## Задание 1

|  |  |  |  |
| --- | --- | --- | --- |
| **№ варианта** | **Исходный формат** | **Результирующий формат** | **День недели** |
| 15 | YAML | JSON | Среда |

Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате YAML. При этом необходимо, чтобы в выбранном дне было не менее двух занятий (можно использовать своё персональное). В случае, если в данный день недели нет таких занятий, то увеличить номер варианта ещё на восемь.

Обязательное задание: написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.

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

Для моего варианта расписание выглядит следующим образом:

Изображение выглядит как текст

Автоматически созданное описание

*Рисунок 1 - Информация о расписании на неделю*

**Содержимое файла на языке YAML:**

---

Schedule:

 table:

  classes:

   lectures:

   labs:

    subject:

     day: Ср

     is\_even\_week: no

     time:

      duration: 11:40-13:10

      weeks:

       - 3

       - 5

       - 7

       - 9

       - 11

       - 13

       - 15

       - 17

     room:

      audience: 2305/1 (бывш. 303/2 (усл)) ауд.

      adress: Кронверкский пр., д.49, лит.А

     lesson:

      name: Информатика

      teacher: Рудникова Тамара Владимировна

      format: Очно - дистанционный

    subject:

     day: Ср

     is\_even\_week: no

     time:

      duration: 13:30-15:00

      weeks:

       - 3

       - 5

       - 7

       - 9

       - 11

       - 13

       - 15

       - 17

     room:

      audience: 2305/1 (бывш. 303/2 (усл)) ауд.

      adress: Кронверкский пр., д.49, лит.А

     lesson:

      name: Информатика

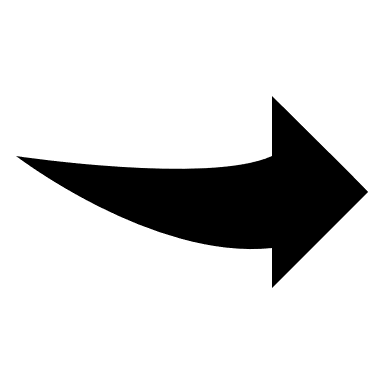
      teacher: Рудникова Тамара Владимировна

      format: Очно - дистанционный

   practices:

   sports:

...

**Код конвертера YAML  JSON без использования сторонних библиотек:**

[https://github.com/ITerNik/ITMOshka/blob/Main/Informatics/Lab\_4/Информатика%20№4(no%20lib).py](https://github.com/ITerNik/ITMOshka/blob/Main/Informatics/Lab_4/%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0%20%E2%84%964(no%20lib).py)

**После конвертации файл выглядит так:**

{

"Schedule": {

"table": {

    "classes": {

        "lectures": null,

            "labs": {

            "subject": {

                "day": "Ср",

                    "is\_even\_week": False,

                    "time": {

                    "duration": "11:40-13:10",

                        "weeks": ['3', '5', '7', '9', '11', '13', '15', '17']

                    },

                    "room": {

                    "audience": "2305/1 (бывш. 303/2 (усл)) ауд.",

                        "adress": "Кронверкский пр., д.49, лит.А"

                    },

                    "lesson": {

                    "name": "Информатика",

                        "teacher": "Рудникова Тамара Владимировна",

                        "format": "Очно - дистанционный"

                    }

                },

                "subject": {

                "day": "Ср",

                    "is\_even\_week": False,

                    "time": {

                    "duration": "13:30-15:00",

                        "weeks": ['3', '5', '7', '9', '11', '13', '15', '17']

                    },

                    "room": {

                    "audience": "2305/1 (бывш. 303/2 (усл)) ауд.",

                        "adress": "Кронверкский пр., д.49, лит.А"

                    },

                    "lesson": {

                    "name": "Информатика",

                        "teacher": "Рудникова Тамара Владимировна",

                        "format": "Очно - дистанционный"

                    }

                }

            },

            "practices": null,

            "sports": null

        }

    }

}

}

## Задание 2

Необходимо было найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.

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

Сравнить полученные результаты и объяснить их сходства и различия.

В интернете большой популярностью пользуется библиотека **PyYaml**, осуществляющая конвертацию YAML файлов в другие форматы

Для ее использование необходимо сначала установить модуль:

Pip install pyyaml

**Со встроенными методами код выглядит намного элегантней:**

[https://github.com/ITerNik/ITMOshka/blob/Main/Informatics/Lab\_4/Информатика%20№4(with%20lib)%20.py](https://github.com/ITerNik/ITMOshka/blob/Main/Informatics/Lab_4/%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0%20%E2%84%964(with%20lib)%20.py)

Для более каноничного вывода JSON используется одноименная библиотека

Судя по всему библиотека использует при конвертации другую кодировку, т.к. символы кириллицы отображаются в файле через Юникод

**Вывод программы:**

{

"Schedule": {

"table": {

    "classes": {

        "lectures": null,

            "labs": {

            "subject": {

                "day": "Ср",

                    "is\_even\_week": False,

                    "time": {

                    "duration": "11:40-13:10",

                        "weeks": ['3', '5', '7', '9', '11', '13', '15', '17']

                    },

                    "room": {

                    "audience": "2305/1 (бывш. 303/2 (усл)) ауд.",

                        "adress": "Кронверкский пр., д.49, лит.А"

                    },

                    "lesson": {

                    "name": "Информатика",

                        "teacher": "Рудникова Тамара Владимировна",

                        "format": "Очно - дистанционный"

                    }

                },

                "subject": {

                "day": "Ср",

                    "is\_even\_week": False,

                    "time": {

                    "duration": "13:30-15:00",

                        "weeks": ['3', '5', '7', '9', '11', '13', '15', '17']

                    },

                    "room": {

                    "audience": "2305/1 (бывш. 303/2 (усл)) ауд.",

                        "adress": "Кронверкский пр., д.49, лит.А"

                    },

                    "lesson": {

                    "name": "Информатика",

                        "teacher": "Рудникова Тамара Владимировна",

                        "format": "Очно - дистанционный"

                    }

                }

            },

            "practices": null,

            "sports": null

        }

    }

}

}

Невооруженным взглядом видна разница в объеме кода двух программ, что не удивительно, ведь большая часть реализации библиотеки скрыта от глаз пользователей

При этом вывод программ отличий практически не имеет. Думаю, мне удалось реализовать парсер на том же уровне функциональности, что и во встроенном модуле

## Задание 3

Необходимо переписать исходный код, добавив в него использование регулярных выражений, а также сравнить полученные результаты

**Код обновленной программы:**

<https://github.com/ITerNik/ITMOshka/blob/Main/Informatics/Lab_4/Информатика%20№4(regex).py>

Регулярные выражения значительно облегчают работу при парсинге входных данных. Логику программы, пусть и с меньшими затратами на оформление корректного вывода, приходится писать с помощью функциональных методов

**Вывод соответствующий:**

{

"Schedule": {

"table": {

    "classes": {

        "lectures": null,

            "labs": {

            "subject": {

                "day": "Ср",

                    "is\_even\_week": False,

                    "time": {

                    "duration": "11:40-13:10",

                       "weeks": ['3', '5', '7', '9', '11', '13', '15', '17']

                    },

                    "room": {

                    "audience": "2305/1 (бывш. 303/2 (усл)) ауд.",

                        "adress": "Кронверкский пр., д.49, лит.А"

                    },

                    "lesson": {

                    "name": "Информатика",

                        "teacher": "Рудникова Тамара Владимировна",

                        "format": "Очно - дистанционный"

                    }

                },

                "subject": {

                "day": "Ср",

                    "is\_even\_week": False,

                    "time": {

                    "duration": "13:30-15:00",

                       "weeks": ['3', '5', '7', '9', '11', '13', '15', '17']

                    },

                    "room": {

                    "audience": "2305/1 (бывш. 303/2 (усл)) ауд.",

                        "adress": "Кронверкский пр., д.49, лит.А"

                    },

                    "lesson": {

                    "name": "Информатика",

                        "teacher": "Рудникова Тамара Владимировна",

                        "format": "Очно - дистанционный"

                    }

                }

            },

            "practices": null,

            "sports": null

        }

    }

}

}

## Задание 4

Используя ранее перечисленные программы из обязательных заданий, сравнить стократное время выполнения парсинга + конвертации в цикле.

Проанализировать полученные результаты.

Время работы программы без библиотек: 0.19960403442382812

Время работы программы со встроенной библиотекой: 0.49784183502197266

Время работы программы с регулярками: 0.19900798797607422

Как видим, встроенная библиотека с огромным отрывом обходит, собственноручно написанные программы, что не удивительно, ввиду ее постоянной поддержки и обновления.

## Вывод

В ходе выполнения данной лабораторной работы мне удалось ближе ознакомиться с языками разметки, включая JSON, YAML, XML, HTML, PROTOBUF, TSV, CSV. Также я вновь попрактиковался в использовании регулярных выражений и сторонних библиотек

## Список литературы.

1. Черновик методического пособия «Информатика»

<https://vk.com/doc-31201840_566998093>

1. Балакшин П.В., Соснин В.В., Калинин И.В., Малышева Т.А., Раков С.В., Рущенко Н.Г., Дергачев А.М. Информатика: лабораторные работы и тесты: Учебно-методическое пособие

<https://books.ifmo.ru/book/2248/informatika:_laboratornye_raboty_i_testy:_uchebno-metodicheskoe_posobie_/_recenzent:_polyakov_v.i..htm>