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

Факультет программной инженерии и компьютерной техники

Лабораторная работа по информатике №4

Исследование протоколов, форматов обмена информацией и языков разметки документов

Вариант №18

Выполнил:  
Студент группы P3106  
Мельник Фёдор Александрович

Проверил:

Балакшин П.В.,

Кандидат технических наук, доцент ФПиКТ

Санкт-Петербург, 2024

# Задание

1. Исходя из структуры расписания конкретного дня, сформировать

файл с расписанием в формате, указанном в задании в качестве

исходного. При этом необходимо, чтобы хотя бы в одной из

выбранных дней 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.

2. Обязательное задание (позволяет набрать до 45 процентов от

максимального числа баллов БаРС за данную лабораторную):

написать программу на языке Python 3.x или любом другом, которая

бы осуществляла парсинг и конвертацию исходного файла в новый

путём простой замены метасимволов исходного формата на

метасимволы результирующего формата.

Нельзя использовать готовые библиотеки, в том числе регулярные

выражения в Python и библиотеки для загрузки XML-файлов.

3. Дополнительное 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. Объяснение должно быть отражено в

отчёте.

4. Дополнительное 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. Объяснение должно быть отражено в

отчёте.

5. Дополнительное задание №3 (позволяет набрать +25 процентов от

максимального числа баллов БаРС за данную лабораторную).

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.

6. Дополнительное задание №4 (позволяет набрать +5 процентов от

максимального числа баллов БаРС за данную лабораторную).

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.

7. Дополнительное задание №5 (позволяет набрать +5 процентов от

максимального числа баллов БаРС за данную лабораторную).

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.

# Решение

## Задание 1

### Исходник JSON

<https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/data/in.json>

{  
 "current\_day": 2,  
 "currant\_week": 9,  
 "date": "Tue, 29 Oct 2024 UTC +03:00",  
 "timetable": [  
 {  
 "title": "Линейная алгебра",  
 "group": "Лин Алг 7.1",  
 "teacher": {  
 "id": 207357,  
 "name": "Тушавин Глеб Владимирович"  
 },  
 "date": {  
 "day": 2,  
 "weeks": [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]  
 },  
 "time": {  
 "lesson\_number": 1,  
 "start": "8:20",  
 "end": "9:50"  
 },  
 "room": {  
 "id": 4210,  
 "name": "ауд. 4210"  
 },  
 "campus": {  
 "id": 0,  
 "name": "ул. Ломоносова, д.9, лит. Б"  
 },  
 "lesson\_type": {  
 "type\_id": 1,  
 "type\_name": "Практика"  
 },  
 "lesson\_format": {  
 "format\_id": 1,  
 "format\_name": "Очно - дистанционный"  
 }  
 },  
 {  
 "title": "Линейная алгебра",  
 "group": "Лин Алг 7",  
 "teacher": {  
 "id": 207357,  
 "name": "Тушавин Глеб Владимирович"  
 },  
 "date": {  
 "day": 2,  
 "weeks": [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]  
 },  
 "time": {  
 "lesson\_number": 2,  
 "start": "10:00",  
 "end": "11:30"  
 },  
 "room": {  
 "id": 2202,  
 "name": "ауд. 2202"  
 },  
 "campus": {  
 "id": 1,  
 "name": "ул. Ломоносова, д.9, лит. A"  
 },  
 "lesson\_type": {  
 "type\_id": 0,  
 "type\_name": "Лекция"  
 },  
 "lesson\_format": {  
 "format\_id": 1,  
 "format\_name": "Очно - дистанционный"  
 }  
 },  
 {  
 "title": "Волейбол",  
 "group": "Волейбол",  
 "teacher": {  
 "id": 269329,  
 "name": "Северюгин Владимир Олегович"  
 },  
 "date": {  
 "day": 2,  
 "weeks": [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]  
 },  
 "time": {  
 "lesson\_number": 3,  
 "start": "11:40",  
 "end": "13:10"  
 },  
 "room": {  
 "id": 4405,  
 "name": "ауд. Большой зал"  
 },  
 "campus": {  
 "id": 0,  
 "name": "ул. Ломоносова, д.9, лит. Б"  
 },  
 "lesson\_type": {  
 "type\_id": 3,  
 "type\_name": "Спорт"  
 },  
 "lesson\_format": {  
 "format\_id": 0,  
 "format\_name": "Очно"  
 }  
 }  
  
 ]  
}

## Задание 2

### Код

main: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/main_task/main.py>

from parser import parse\_json  
from parser.converter import convert\_to\_xml  
from os import path  
  
json = open(path.join(path.dirname(\_\_file\_\_), "../data/in.json"), 'r').read()  
xml = open(path.join(path.dirname(\_\_file\_\_), "../data/out.xml"), 'w')  
xml.write('<?xml version="1.0" encoding="UTF-8" ?>' + convert\_to\_xml(parse\_json(json)))

parser: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/main_task/parser/__init__.py>

def parse\_json(json: str):  
 string = parse\_value(json.strip())[0]  
 return string  
  
  
def parse\_keyvalue(string: str):  
 parsed\_string = parse\_string(string)  
 if parsed\_string is None:  
 return None  
 parsed\_colon = parse\_colon(parsed\_string[1])  
 if parsed\_colon is None:  
 return None  
 parsed\_value = parse\_value(parsed\_colon[1])  
 if parsed\_value is None:  
 return None  
 return (parsed\_string[0], parsed\_value[0]), parsed\_value[1].strip()  
  
  
def parse\_value(string: str):  
 parsed\_number = parse\_number(string)  
 if not parsed\_number is None:  
 parse\_value(parsed\_number[1].strip())  
 return parsed\_number[0], parsed\_number[1].strip()  
 parsed\_string = parse\_string(string)  
 if not parsed\_string is None:  
 parse\_value(parsed\_string[1].strip())  
 return parsed\_string[0], parsed\_string[1].strip()  
 parsed\_array = parse\_array(string)  
 if not parsed\_array is None:  
 parse\_value(parsed\_array[1].strip())  
 return parsed\_array[0], parsed\_array[1].strip()  
 parsed\_obect = parse\_object(string)  
 if not parsed\_obect is None:  
 parse\_value(parsed\_obect[1].strip())  
 return parsed\_obect[0], parsed\_obect[1].strip()  
  
  
def parse\_comma(string: str):  
 if not string.startswith(','):  
 return None  
 return ",", string[1:].strip()  
  
  
def parse\_comma\_values(string: str):  
 res = []  
 while True:  
 parsed\_value = parse\_value(string)  
 if parsed\_value is None:  
 break  
 res.append(parsed\_value[0])  
 string = parsed\_value[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1]  
 return res, string.strip()  
  
  
def parse\_array(string: str):  
 if not string.startswith('['):  
 return None  
 parsed\_comma\_values = parse\_comma\_values(string[1:].strip())  
 if parsed\_comma\_values is None:  
 res, string = [], string[1:]  
 else:  
 res, string = parsed\_comma\_values  
 if not string.startswith(']'):  
 return None  
 return res, string[1:].strip()  
  
  
def parse\_string(string: str):  
 if not string.startswith("\""):  
 return None  
 second\_quote = string.find("\"", 1)  
 if second\_quote == -1:  
 return None  
 return string[1:second\_quote], string[second\_quote + 1:]  
  
  
def parse\_colon(string: str):  
 if not string.startswith(":"):  
 return None  
 else:  
 return ":", string[2:].strip()  
  
  
def parse\_comma\_keyvalues(string: str):  
 res = {}  
 while True:  
 parsed\_keyvalue = parse\_keyvalue(string)  
 if parsed\_keyvalue is None:  
 break  
 key, val = parsed\_keyvalue[0]  
 res[key] = val  
 string = parsed\_keyvalue[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1].strip()  
 return res, string.strip()  
  
  
def parse\_object(string: str):  
 if not string.startswith("{"):  
 return None  
 arr, string = parse\_comma\_keyvalues(string[1:].strip())  
 if not string.startswith("}"):  
 return None  
 return arr, string[1:].strip()  
  
  
def parse\_number(string: str):  
 num = ""  
 for i in string:  
 if i in "0123456789.Ee-+":  
 num += i  
 else:  
 break  
 if num.isnumeric():  
 return int(num), string[len(num):].strip()  
 else:  
 try:  
 return float(num), string[len(num):].strip()  
 except ValueError:  
 return None

converter: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/main_task/parser/converter/__init__.py>

def convert\_to\_xml(d: dict, tag="root"):  
 xml = ""  
 for key, value in d.items():  
 if isinstance(value, dict):  
 xml += convert\_to\_xml(value, tag=key)  
 elif isinstance(value, list):  
 xml += f"<{key}>"  
 for i in value:  
 xml += "".join(convert\_to\_xml({key + "\_element": i}, tag=""))  
 xml += f"</{key}>"  
 else:  
 xml += f"<{key}>" + str(value) + f"</{key}>"  
 if tag == "":  
 return xml  
 else:  
 return f"<{tag}>" + xml + f"</{tag}>"

### Результат

<https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/data/out.xml>

<?xml version="1.0" encoding="UTF-8" ?><root><current\_day>2</current\_day><currant\_week>9</currant\_week><date>Tue, 29 Oct 2024 UTC +03:00</date><timetable><timetable\_element><title>Линейная алгебра</title><group>Лин Алг 7.1</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>1</lesson\_number><start>8:20</start><end>9:50</end></time><room><id>4210</id><name>ауд. 4210</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>1</type\_id><type\_name>Практика</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>Линейная алгебра</title><group>Лин Алг 7</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>2</lesson\_number><start>10:00</start><end>11:30</end></time><room><id>2202</id><name>ауд. 2202</name></room><campus><id>1</id><name>ул. Ломоносова, д.9, лит. A</name></campus><lesson\_type><type\_id>0</type\_id><type\_name>Лекция</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>Волейбол</title><group>Волейбол</group><teacher><id>269329</id><name>Северюгин Владимир Олегович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>3</lesson\_number><start>11:40</start><end>13:10</end></time><room><id>4405</id><name>ауд. Большой зал</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>3</type\_id><type\_name>Спорт</type\_name></lesson\_type><lesson\_format><format\_id>0</format\_id><format\_name>Очно</format\_name></lesson\_format></timetable\_element></timetable></root>

Читаемый формат:

<?xml version="1.0" encoding="UTF-8" ?>  
<root>  
 <current\_day>2</current\_day>  
 <currant\_week>9</currant\_week>  
 <date>Tue, 29 Oct 2024 UTC +03:00</date>  
 <timetable>  
 <timetable\_element>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7.1</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>1</lesson\_number>  
 <start>8:20</start>  
 <end>9:50</end>  
 </time>  
 <room>  
 <id>4210</id>  
 <name>ауд. 4210</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>1</type\_id>  
 <type\_name>Практика</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>2</lesson\_number>  
 <start>10:00</start>  
 <end>11:30</end>  
 </time>  
 <room>  
 <id>2202</id>  
 <name>ауд. 2202</name>  
 </room>  
 <campus>  
 <id>1</id>  
 <name>ул. Ломоносова, д.9, лит. A</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>0</type\_id>  
 <type\_name>Лекция</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>Волейбол</title>  
 <group>Волейбол</group>  
 <teacher>  
 <id>269329</id>  
 <name>Северюгин Владимир Олегович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>3</lesson\_number>  
 <start>11:40</start>  
 <end>13:10</end>  
 </time>  
 <room>  
 <id>4405</id>  
 <name>ауд. Большой зал</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>3</type\_id>  
 <type\_name>Спорт</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>0</format\_id>  
 <format\_name>Очно</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 </timetable>  
</root>

## Задание 3

### Код

main: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/additional1/main.py>

from json import loads  
from os import path  
from dicttoxml import dicttoxml  
  
  
def convert\_to\_xml(json\_data):  
 xml = dicttoxml(json\_data, attr\_type=False, return\_bytes=False)  
 return xml  
  
  
def parse\_json(string: str):  
 json\_data = loads(string)  
 return json\_data  
  
  
if \_\_name\_\_ == "\_\_main\_\_":  
 json = open(path.join(path.dirname(\_\_file\_\_), "../data/in.json"), 'r').read()  
 xml = open(path.join(path.dirname(\_\_file\_\_), "../data/out-add1.xml"), 'w')  
 xml.write(convert\_to\_xml(parse\_json(json)))

### Результат

<https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/data/out-add1.xml>

<?xml version="1.0" encoding="UTF-8" ?><root><current\_day>2</current\_day><currant\_week>9</currant\_week><date>Tue, 29 Oct 2024 UTC +03:00</date><timetable><item><title>Линейная алгебра</title><group>Лин Алг 7.1</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><item>2</item><item>3</item><item>4</item><item>5</item><item>6</item><item>7</item><item>8</item><item>9</item><item>10</item><item>11</item><item>12</item><item>13</item><item>14</item><item>15</item><item>16</item><item>17</item></weeks></date><time><lesson\_number>1</lesson\_number><start>8:20</start><end>9:50</end></time><room><id>4210</id><name>ауд. 4210</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>1</type\_id><type\_name>Практика</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></item><item><title>Линейная алгебра</title><group>Лин Алг 7</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><item>2</item><item>3</item><item>4</item><item>5</item><item>6</item><item>7</item><item>8</item><item>9</item><item>10</item><item>11</item><item>12</item><item>13</item><item>14</item><item>15</item><item>16</item><item>17</item></weeks></date><time><lesson\_number>2</lesson\_number><start>10:00</start><end>11:30</end></time><room><id>2202</id><name>ауд. 2202</name></room><campus><id>1</id><name>ул. Ломоносова, д.9, лит. A</name></campus><lesson\_type><type\_id>0</type\_id><type\_name>Лекция</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></item><item><title>Волейбол</title><group>Волейбол</group><teacher><id>269329</id><name>Северюгин Владимир Олегович</name></teacher><date><day>2</day><weeks><item>2</item><item>3</item><item>4</item><item>5</item><item>6</item><item>7</item><item>8</item><item>9</item><item>10</item><item>11</item><item>12</item><item>13</item><item>14</item><item>15</item><item>16</item><item>17</item></weeks></date><time><lesson\_number>3</lesson\_number><start>11:40</start><end>13:10</end></time><room><id>4405</id><name>ауд. Большой зал</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>3</type\_id><type\_name>Спорт</type\_name></lesson\_type><lesson\_format><format\_id>0</format\_id><format\_name>Очно</format\_name></lesson\_format></item></timetable></root>

Читаемый формат:

<?xml version="1.0" encoding="UTF-8" ?>  
<root>  
 <current\_day>2</current\_day>  
 <currant\_week>9</currant\_week>  
 <date>Tue, 29 Oct 2024 UTC +03:00</date>  
 <timetable>  
 <item>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7.1</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <item>2</item>  
 <item>3</item>  
 <item>4</item>  
 <item>5</item>  
 <item>6</item>  
 <item>7</item>  
 <item>8</item>  
 <item>9</item>  
 <item>10</item>  
 <item>11</item>  
 <item>12</item>  
 <item>13</item>  
 <item>14</item>  
 <item>15</item>  
 <item>16</item>  
 <item>17</item>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>1</lesson\_number>  
 <start>8:20</start>  
 <end>9:50</end>  
 </time>  
 <room>  
 <id>4210</id>  
 <name>ауд. 4210</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>1</type\_id>  
 <type\_name>Практика</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </item>  
 <item>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <item>2</item>  
 <item>3</item>  
 <item>4</item>  
 <item>5</item>  
 <item>6</item>  
 <item>7</item>  
 <item>8</item>  
 <item>9</item>  
 <item>10</item>  
 <item>11</item>  
 <item>12</item>  
 <item>13</item>  
 <item>14</item>  
 <item>15</item>  
 <item>16</item>  
 <item>17</item>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>2</lesson\_number>  
 <start>10:00</start>  
 <end>11:30</end>  
 </time>  
 <room>  
 <id>2202</id>  
 <name>ауд. 2202</name>  
 </room>  
 <campus>  
 <id>1</id>  
 <name>ул. Ломоносова, д.9, лит. A</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>0</type\_id>  
 <type\_name>Лекция</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </item>  
 <item>  
 <title>Волейбол</title>  
 <group>Волейбол</group>  
 <teacher>  
 <id>269329</id>  
 <name>Северюгин Владимир Олегович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <item>2</item>  
 <item>3</item>  
 <item>4</item>  
 <item>5</item>  
 <item>6</item>  
 <item>7</item>  
 <item>8</item>  
 <item>9</item>  
 <item>10</item>  
 <item>11</item>  
 <item>12</item>  
 <item>13</item>  
 <item>14</item>  
 <item>15</item>  
 <item>16</item>  
 <item>17</item>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>3</lesson\_number>  
 <start>11:40</start>  
 <end>13:10</end>  
 </time>  
 <room>  
 <id>4405</id>  
 <name>ауд. Большой зал</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>3</type\_id>  
 <type\_name>Спорт</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>0</format\_id>  
 <format\_name>Очно</format\_name>  
 </lesson\_format>  
 </item>  
 </timetable>  
</root>

### Сравнение результатов [задания 2](#_Результат) и [задания 3](#_Результат_1)

Результаты отличаются форматом вывода элементов массива. Если в случае задания 2 тэги элемента массива назывались <(list\_name)\_element>, то в задании 3 теги элемента массива называются <item>. Это происходит из-за того, что во втором задании мы сами выбрали название тэгов, а в третьем библиотека сама поставила тегу название по умолчанию.

## Задание 4

### Код

main: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/additional2/main.py>

from parser import parse\_json  
from parser.converter import convert\_to\_xml  
from os import path  
  
json = open(path.join(path.dirname(\_\_file\_\_), "../data/in.json"), 'r').read()  
xml = open(path.join(path.dirname(\_\_file\_\_), "../data/out-add2.xml"), 'w')  
xml.write('<?xml version="1.0" encoding="UTF-8" ?>' + convert\_to\_xml(parse\_json(json)))

parser: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/additional2/parser/__init__.py>

import re  
from re import compile  
  
string\_regex = compile(r"(\"(?:[^\\\"]|\\[\"nfbtr])\*?\")(\s\*.\*)", re.DOTALL)  
number\_regex = compile(r"(-?(?:0|[1-9]\d\*)(?:\.\d\*)?(?:[eE][+-]?\d+)?)(\s\*.\*)", re.DOTALL)  
null\_regex = compile(r"(null)\s\*(.\*)", re.DOTALL)  
bool\_regex = compile(r"(true|false)\s\*(.\*)", re.DOTALL)  
  
  
def parse\_json(json: str):  
 string = parse\_value(json.strip())[0]  
 return string  
  
  
def parse\_keyvalue(string: str):  
 parsed\_string = parse\_string(string)  
 if parsed\_string is None:  
 return None  
 parsed\_colon = parse\_colon(parsed\_string[1])  
 if parsed\_colon is None:  
 return None  
 parsed\_value = parse\_value(parsed\_colon[1])  
 if parsed\_value is None:  
 return None  
 return (parsed\_string[0], parsed\_value[0]), parsed\_value[1].strip()  
  
  
def parse\_value(string: str):  
 parsed\_number = parse\_number(string)  
 if not parsed\_number is None:  
 parse\_value(parsed\_number[1].strip())  
 return parsed\_number[0], parsed\_number[1].strip()  
 parsed\_string = parse\_string(string)  
 if not parsed\_string is None:  
 parse\_value(parsed\_string[1].strip())  
 return parsed\_string[0], parsed\_string[1].strip()  
 parsed\_bool = parse\_bool(string)  
 if not parsed\_bool is None:  
 parse\_value(parsed\_bool[1].strip())  
 return parsed\_bool[0], parsed\_bool[1].strip()  
 parsed\_null = parse\_null(string)  
 if not parsed\_null is None:  
 parse\_value(parsed\_null[1].strip())  
 return parsed\_null[0], parsed\_null[1].strip()  
 parsed\_array = parse\_array(string)  
 if not parsed\_array is None:  
 parse\_value(parsed\_array[1].strip())  
 return parsed\_array[0], parsed\_array[1].strip()  
 parsed\_obect = parse\_object(string)  
 if not parsed\_obect is None:  
 parse\_value(parsed\_obect[1].strip())  
 return parsed\_obect[0], parsed\_obect[1].strip()  
  
  
def parse\_bool(string: str):  
 bool\_match = bool\_regex.match(string.strip())  
 if not bool\_match:  
 return None  
 res, string = bool\_match.groups()  
 return res == 'true', string.strip()  
  
  
def parse\_null(string: str):  
 null\_match = null\_regex.match(string.strip())  
 if not null\_match:  
 return None  
 res, string = null\_match.groups()  
 return 'null', string.strip()  
  
  
def parse\_comma(string: str):  
 if not string.startswith(','):  
 return None  
 return ",", string[1:].strip()  
  
  
def parse\_comma\_values(string: str):  
 res = []  
 while True:  
 parsed\_value = parse\_value(string)  
 if parsed\_value is None:  
 break  
 res.append(parsed\_value[0])  
 string = parsed\_value[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1]  
 return res, string.strip()  
  
  
def parse\_array(string: str):  
 if not string.startswith('['):  
 return None  
 parsed\_comma\_values = parse\_comma\_values(string[1:].strip())  
 if parsed\_comma\_values is None:  
 res, string = [], string[1:]  
 else:  
 res, string = parsed\_comma\_values  
 if not string.startswith(']'):  
 return None  
 return res, string[1:].strip()  
  
  
def parse\_string(string: str):  
 string\_match = string\_regex.match(string.strip())  
 if not string\_match:  
 return None  
 res, string = string\_match.groups()  
 return res[1:-1], string.strip()  
  
  
def parse\_colon(string: str):  
 if not string.startswith(":"):  
 return None  
 else:  
 return ":", string[2:].strip()  
  
  
def parse\_comma\_keyvalues(string: str):  
 res = {}  
 while True:  
 parsed\_keyvalue = parse\_keyvalue(string)  
 if parsed\_keyvalue is None:  
 break  
 key, val = parsed\_keyvalue[0]  
 res[key] = val  
 string = parsed\_keyvalue[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1].strip()  
 return res, string.strip()  
  
  
def parse\_object(string: str):  
 if not string.startswith("{"):  
 return None  
 arr, string = parse\_comma\_keyvalues(string[1:].strip())  
 if not string.startswith("}"):  
 return None  
 return arr, string[1:].strip()  
  
  
def parse\_number(string: str):  
 number\_match = number\_regex.match(string.strip())  
 if not number\_match:  
 return None  
 res, string = number\_match.groups()  
 return eval(res), string.strip()

converter: <https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/additional2/parser/converter/__init__.py>

def convert\_to\_xml(d: dict, tag="root"):  
 xml = ""  
 for key, value in d.items():  
 if isinstance(value, dict):  
 xml += convert\_to\_xml(value, tag=key)  
 elif isinstance(value, list):  
 xml += f"<{key}>"  
 for i in value:  
 xml += "".join(convert\_to\_xml({key + "\_element": i}, tag=""))  
 xml += f"</{key}>"  
 else:  
 xml += f"<{key}>" + str(value) + f"</{key}>"  
 if tag == "":  
 return xml  
 else:  
 return f"<{tag}>" + xml + f"</{tag}>"

### Результат

<https://github.com/ldpst/itmo/blob/main/labs/sem1/inf/lab4/tasks/data/out-add2.xml>

<?xml version="1.0" encoding="UTF-8" ?><root><current\_day>2</current\_day><currant\_week>9</currant\_week><date>Tue, 29 Oct 2024 UTC +03:00</date><timetable><timetable\_element><title>Линейная алгебра</title><group>Лин Алг 7.1</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>1</lesson\_number><start>8:20</start><end>9:50</end></time><room><id>4210</id><name>ауд. 4210</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>1</type\_id><type\_name>Практика</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>Линейная алгебра</title><group>Лин Алг 7</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>2</lesson\_number><start>10:00</start><end>11:30</end></time><room><id>2202</id><name>ауд. 2202</name></room><campus><id>1</id><name>ул. Ломоносова, д.9, лит. A</name></campus><lesson\_type><type\_id>0</type\_id><type\_name>Лекция</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>Волейбол</title><group>Волейбол</group><teacher><id>269329</id><name>Северюгин Владимир Олегович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>3</lesson\_number><start>11:40</start><end>13:10</end></time><room><id>4405</id><name>ауд. Большой зал</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>3</type\_id><type\_name>Спорт</type\_name></lesson\_type><lesson\_format><format\_id>0</format\_id><format\_name>Очно</format\_name></lesson\_format></timetable\_element></timetable></root>

В читаемом формате:

<?xml version="1.0" encoding="UTF-8" ?>  
<root>  
 <current\_day>2</current\_day>  
 <currant\_week>9</currant\_week>  
 <date>Tue, 29 Oct 2024 UTC +03:00</date>  
 <timetable>  
 <timetable\_element>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7.1</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>1</lesson\_number>  
 <start>8:20</start>  
 <end>9:50</end>  
 </time>  
 <room>  
 <id>4210</id>  
 <name>ауд. 4210</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>1</type\_id>  
 <type\_name>Практика</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>2</lesson\_number>  
 <start>10:00</start>  
 <end>11:30</end>  
 </time>  
 <room>  
 <id>2202</id>  
 <name>ауд. 2202</name>  
 </room>  
 <campus>  
 <id>1</id>  
 <name>ул. Ломоносова, д.9, лит. A</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>0</type\_id>  
 <type\_name>Лекция</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>Волейбол</title>  
 <group>Волейбол</group>  
 <teacher>  
 <id>269329</id>  
 <name>Северюгин Владимир Олегович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>3</lesson\_number>  
 <start>11:40</start>  
 <end>13:10</end>  
 </time>  
 <room>  
 <id>4405</id>  
 <name>ауд. Большой зал</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>3</type\_id>  
 <type\_name>Спорт</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>0</format\_id>  
 <format\_name>Очно</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 </timetable>  
</root>

### Сравнение результатов [задания 2](#_Результат_2) и [задания 4](#_Результат_3)

Результаты полностью совпадают, так как изменен лишь принцип парсинга JSON файла, а конвертация не изменялась.

## Задание 5

### Исходник JSON

<https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/data/in-add3.json>

{  
 "current\_day": 2,  
 "currant\_week": 9,  
 "date": "Tue, 29 Oct 2024 UTC +03:00",  
 "timetable": [  
 {  
 "title": "Линейная алгебра",  
 "group": "Лин Алг 7.1",  
 "teacher": {  
 "id": 207357,  
 "name": "Тушавин Глеб Владимирович"  
 },  
 "date": {  
 "day": 2,  
 "weeks": [  
 2,  
 3,  
 4,  
 5,  
 6,  
 7,  
 8,  
 9,  
 10,  
 11,  
 12,  
 13,  
 14,  
 15,  
 16,  
 17  
 ]  
 },  
 "time": {  
 "lesson\_number": 1,  
 "start": "8:20",  
 "end": "9:50"  
 },  
 "room": {  
 "id": 4210,  
 "name": "ауд. 4210"  
 },  
 "campus": {  
 "id": 0,  
 "name": "ул. Ломоносова, д.9, лит. Б"  
 },  
 "lesson\_type": {  
 "type\_id": 1,  
 "type\_name": "Практика"  
 },  
 "lesson\_format": {  
 "format\_id": 1,  
 "format\_name": "Очно - дистанционный"  
 }  
 },  
 {  
 "title": "Линейная алгебра",  
 "group": "Лин Алг 7",  
 "teacher": {  
 "id": 207357,  
 "name": "Тушавин Глеб Владимирович"  
 },  
 "date": {  
 "day": 2,  
 "weeks": [  
 2,  
 3,  
 4,  
 5,  
 6,  
 7,  
 8,  
 9,  
 10,  
 11,  
 12,  
 13,  
 14,  
 15,  
 16,  
 17  
 ]  
 },  
 "time": {  
 "lesson\_number": 2,  
 "start": "10:00",  
 "end": "11:30"  
 },  
 "room": {  
 "id": 2202,  
 "name": "ауд. 2202"  
 },  
 "campus": {  
 "id": 1,  
 "name": "ул. Ломоносова, д.9, лит. A"  
 },  
 "lesson\_type": {  
 "type\_id": 0,  
 "type\_name": "Лекция"  
 },  
 "lesson\_format": {  
 "format\_id": 1,  
 "format\_name": "Очно - дистанционный"  
 }  
 },  
 {  
 "title": "Волейбол",  
 "group": "Волейбол",  
 "teacher": {  
 "id": 269329,  
 "name": "Северюгин Владимир Олегович"  
 },  
 "date": {  
 "day": 2,  
 "weeks": [  
 2,  
 3,  
 4,  
 5,  
 6,  
 7,  
 8,  
 9,  
 10,  
 11,  
 12,  
 13,  
 14,  
 15,  
 16,  
 17  
 ]  
 },  
 "time": {  
 "lesson\_number": 3,  
 "start": "11:40",  
 "end": "13:10"  
 },  
 "room": {  
 "id": 4405,  
 "name": "ауд. Большой зал"  
 },  
 "campus": {  
 "id": 0,  
 "name": "ул. Ломоносова, д.9, лит. Б"  
 },  
 "lesson\_type": {  
 "type\_id": 3,  
 "type\_name": "Спорт"  
 },  
 "lesson\_format": {  
 "format\_id": 0,  
 "format\_name": "Очно"  
 }  
 },  
 {  
 "title": "История российской науки и техники",  
 "group": "ИНТ 2",  
 "teacher": {  
 "id": 340852,  
 "name": "Белоусов Александр Сергеевич"  
 },  
 "date": {  
 "day": 4,  
 "weeks": [  
 1,  
 3,  
 5,  
 7,  
 9,  
 13,  
 15  
 ]  
 },  
 "time": {  
 "lesson\_number": 4,  
 "start": "13:30",  
 "end": "15:00"  
 },  
 "room": {  
 "id": 1419,  
 "name": "ауд. Lemon Classroom"  
 },  
 "campus": {  
 "id": 2,  
 "name": "Кронверкский пр., д.49, лит.А"  
 },  
 "lesson\_type": {  
 "type\_id": 0,  
 "type\_name": "Лекция"  
 },  
 "lesson\_format": {  
 "format\_id": 1,  
 "format\_name": "Очно - дистанционный"  
 }  
 },  
 {  
 "title": "История российской науки и техники",  
 "group": "ИНТ 2.2",  
 "teacher": {  
 "id": 340852,  
 "name": "Белоусов Александр Сергеевич"  
 },  
 "date": {  
 "day": 4,  
 "weeks": [  
 1,  
 3,  
 5,  
 7,  
 9,  
 11,  
 13,  
 15  
 ]  
 },  
 "time": {  
 "lesson\_number": 5,  
 "start": "15:20",  
 "end": "16:50"  
 },  
 "room": {  
 "id": 2414,  
 "name": "ауд. 2414"  
 },  
 "campus": {  
 "id": 2,  
 "name": "Кронверкский пр., д.49, лит.А"  
 },  
 "lesson\_type": {  
 "type\_id": 1,  
 "type\_name": "Практика"  
 },  
 "lesson\_format": {  
 "format\_id": 1,  
 "format\_name": "Очно - дистанционный"  
 }  
 }  
 ]  
}

### Код

main: <https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/additional3/main.py>

from parser import parse\_json  
from parser.converter import convert\_to\_xml  
from os import path  
  
json = open(path.join(path.dirname(\_\_file\_\_), "../data/in-add3.json"), 'r').read()  
xml = open(path.join(path.dirname(\_\_file\_\_), "../data/out-add3.xml"), 'w')  
xml.write('<?xml version="1.0" encoding="UTF-8" ?>' + convert\_to\_xml(parse\_json(json)))

parser: <https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/additional3/parser/__init__.py>

def parse\_json(json: str):  
 string = parse\_value(json.strip())[0]  
 return string  
  
  
def parse\_keyvalue(string: str):  
 parsed\_string = parse\_string(string)  
 if parsed\_string is None:  
 return None  
 parsed\_colon = parse\_colon(parsed\_string[1])  
 if parsed\_colon is None:  
 return None  
 parsed\_value = parse\_value(parsed\_colon[1])  
 if parsed\_value is None:  
 return None  
 return (parsed\_string[0], parsed\_value[0]), parsed\_value[1].strip()  
  
  
def parse\_value(string: str):  
 parsed\_number = parse\_number(string)  
 if not parsed\_number is None:  
 parse\_value(parsed\_number[1].strip())  
 return parsed\_number[0], parsed\_number[1].strip()  
 parsed\_string = parse\_string(string)  
 if not parsed\_string is None:  
 parse\_value(parsed\_string[1].strip())  
 return parsed\_string[0], parsed\_string[1].strip()  
 parsed\_bool = parse\_bool(string)  
 if not parsed\_bool is None:  
 parse\_value(parsed\_bool[1].strip())  
 return parsed\_bool[0], parsed\_bool[1].strip()  
 parsed\_null = parse\_null(string)  
 if not parsed\_null is None:  
 parse\_value(parsed\_null[1].strip())  
 return parsed\_null[0], parsed\_null[1].strip()  
 parsed\_array = parse\_array(string)  
 if not parsed\_array is None:  
 parse\_value(parsed\_array[1].strip())  
 return parsed\_array[0], parsed\_array[1].strip()  
 parsed\_obect = parse\_object(string)  
 if not parsed\_obect is None:  
 parse\_value(parsed\_obect[1].strip())  
 return parsed\_obect[0], parsed\_obect[1].strip()  
  
  
def parse\_bool(string: str):  
 if string.startswith("true"):  
 return True, string[4:].strip()  
 if string.startswith("false"):  
 return False, string[5:].strip()  
  
  
def parse\_null(string: str):  
 if string.startswith("null"):  
 return None, string[4:].strip()  
  
  
def parse\_comma(string: str):  
 if not string.startswith(','):  
 return None  
 return ",", string[1:].strip()  
  
  
def parse\_comma\_values(string: str):  
 res = []  
 while True:  
 parsed\_value = parse\_value(string)  
 if parsed\_value is None:  
 break  
 res.append(parsed\_value[0])  
 string = parsed\_value[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1]  
 return res, string.strip()  
  
  
def parse\_array(string: str):  
 if not string.startswith('['):  
 return None  
 parsed\_comma\_values = parse\_comma\_values(string[1:].strip())  
 if parsed\_comma\_values is None:  
 res, string = [], string[1:]  
 else:  
 res, string = parsed\_comma\_values  
 if not string.startswith(']'):  
 return None  
 return res, string[1:].strip()  
  
  
def parse\_string(string: str):  
 if not string.startswith("\""):  
 return None  
 second\_quote = string.find("\"", 1)  
 if second\_quote == -1:  
 return None  
 return string[1:second\_quote], string[second\_quote + 1:]  
  
  
def parse\_colon(string: str):  
 if not string.startswith(":"):  
 return None  
 else:  
 return ":", string[2:].strip()  
  
  
def parse\_comma\_keyvalues(string: str):  
 res = {}  
 while True:  
 parsed\_keyvalue = parse\_keyvalue(string)  
 if parsed\_keyvalue is None:  
 break  
 key, val = parsed\_keyvalue[0]  
 res[key] = val  
 string = parsed\_keyvalue[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1].strip()  
 return res, string.strip()  
  
  
def parse\_object(string: str):  
 if not string.startswith("{"):  
 return None  
 arr, string = parse\_comma\_keyvalues(string[1:].strip())  
 if not string.startswith("}"):  
 return None  
 return arr, string[1:].strip()  
  
  
def parse\_number(string: str):  
 num = ""  
 for i in string:  
 if i in "0123456789.Ee-+":  
 num += i  
 else:  
 break  
 if num.isnumeric():  
 return int(num), string[len(num):].strip()  
 else:  
 try:  
 return float(num), string[len(num):].strip()  
 except ValueError:  
 return None

converter: <https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/additional3/parser/converter/__init__.py>

def convert\_to\_xml(d: dict, tag="root"):  
 xml = ""  
 for key, value in d.items():  
 if isinstance(value, dict):  
 xml += convert\_to\_xml(value, tag=key)  
 elif isinstance(value, list):  
 xml += f"<{key}>"  
 for i in value:  
 xml += "".join(convert\_to\_xml({key + "\_element": i}, tag=""))  
 xml += f"</{key}>"  
 else:  
 xml += f"<{key}>" + str(value) + f"</{key}>"  
 if tag == "":  
 return xml  
 else:  
 return f"<{tag}>" + xml + f"</{tag}>"

### Результат

<https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/data/out-add3.xml>

<?xml version="1.0" encoding="UTF-8" ?><root><current\_day>2</current\_day><currant\_week>9</currant\_week><date>Tue, 29 Oct 2024 UTC +03:00</date><timetable><timetable\_element><title>Линейная алгебра</title><group>Лин Алг 7.1</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>1</lesson\_number><start>8:20</start><end>9:50</end></time><room><id>4210</id><name>ауд. 4210</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>1</type\_id><type\_name>Практика</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>Линейная алгебра</title><group>Лин Алг 7</group><teacher><id>207357</id><name>Тушавин Глеб Владимирович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>2</lesson\_number><start>10:00</start><end>11:30</end></time><room><id>2202</id><name>ауд. 2202</name></room><campus><id>1</id><name>ул. Ломоносова, д.9, лит. A</name></campus><lesson\_type><type\_id>0</type\_id><type\_name>Лекция</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>Волейбол</title><group>Волейбол</group><teacher><id>269329</id><name>Северюгин Владимир Олегович</name></teacher><date><day>2</day><weeks><weeks\_element>2</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>4</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>6</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>8</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>10</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>12</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>14</weeks\_element><weeks\_element>15</weeks\_element><weeks\_element>16</weeks\_element><weeks\_element>17</weeks\_element></weeks></date><time><lesson\_number>3</lesson\_number><start>11:40</start><end>13:10</end></time><room><id>4405</id><name>ауд. Большой зал</name></room><campus><id>0</id><name>ул. Ломоносова, д.9, лит. Б</name></campus><lesson\_type><type\_id>3</type\_id><type\_name>Спорт</type\_name></lesson\_type><lesson\_format><format\_id>0</format\_id><format\_name>Очно</format\_name></lesson\_format></timetable\_element><timetable\_element><title>История российской науки и техники</title><group>ИНТ 2</group><teacher><id>340852</id><name>Белоусов Александр Сергеевич</name></teacher><date><day>4</day><weeks><weeks\_element>1</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>15</weeks\_element></weeks></date><time><lesson\_number>4</lesson\_number><start>13:30</start><end>15:00</end></time><room><id>1419</id><name>ауд. Lemon Classroom</name></room><campus><id>2</id><name>Кронверкский пр., д.49, лит.А</name></campus><lesson\_type><type\_id>0</type\_id><type\_name>Лекция</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element><timetable\_element><title>История российской науки и техники</title><group>ИНТ 2.2</group><teacher><id>340852</id><name>Белоусов Александр Сергеевич</name></teacher><date><day>4</day><weeks><weeks\_element>1</weeks\_element><weeks\_element>3</weeks\_element><weeks\_element>5</weeks\_element><weeks\_element>7</weeks\_element><weeks\_element>9</weeks\_element><weeks\_element>11</weeks\_element><weeks\_element>13</weeks\_element><weeks\_element>15</weeks\_element></weeks></date><time><lesson\_number>5</lesson\_number><start>15:20</start><end>16:50</end></time><room><id>2414</id><name>ауд. 2414</name></room><campus><id>2</id><name>Кронверкский пр., д.49, лит.А</name></campus><lesson\_type><type\_id>1</type\_id><type\_name>Практика</type\_name></lesson\_type><lesson\_format><format\_id>1</format\_id><format\_name>Очно - дистанционный</format\_name></lesson\_format></timetable\_element></timetable></root>

В читаемом виде:

<?xml version="1.0" encoding="UTF-8" ?>  
<root>  
 <current\_day>2</current\_day>  
 <currant\_week>9</currant\_week>  
 <date>Tue, 29 Oct 2024 UTC +03:00</date>  
 <timetable>  
 <timetable\_element>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7.1</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>1</lesson\_number>  
 <start>8:20</start>  
 <end>9:50</end>  
 </time>  
 <room>  
 <id>4210</id>  
 <name>ауд. 4210</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>1</type\_id>  
 <type\_name>Практика</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>Линейная алгебра</title>  
 <group>Лин Алг 7</group>  
 <teacher>  
 <id>207357</id>  
 <name>Тушавин Глеб Владимирович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>2</lesson\_number>  
 <start>10:00</start>  
 <end>11:30</end>  
 </time>  
 <room>  
 <id>2202</id>  
 <name>ауд. 2202</name>  
 </room>  
 <campus>  
 <id>1</id>  
 <name>ул. Ломоносова, д.9, лит. A</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>0</type\_id>  
 <type\_name>Лекция</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>Волейбол</title>  
 <group>Волейбол</group>  
 <teacher>  
 <id>269329</id>  
 <name>Северюгин Владимир Олегович</name>  
 </teacher>  
 <date>  
 <day>2</day>  
 <weeks>  
 <weeks\_element>2</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>4</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>6</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>8</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>10</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>12</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>14</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 <weeks\_element>16</weeks\_element>  
 <weeks\_element>17</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>3</lesson\_number>  
 <start>11:40</start>  
 <end>13:10</end>  
 </time>  
 <room>  
 <id>4405</id>  
 <name>ауд. Большой зал</name>  
 </room>  
 <campus>  
 <id>0</id>  
 <name>ул. Ломоносова, д.9, лит. Б</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>3</type\_id>  
 <type\_name>Спорт</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>0</format\_id>  
 <format\_name>Очно</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>История российской науки и техники</title>  
 <group>ИНТ 2</group>  
 <teacher>  
 <id>340852</id>  
 <name>Белоусов Александр Сергеевич</name>  
 </teacher>  
 <date>  
 <day>4</day>  
 <weeks>  
 <weeks\_element>1</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>4</lesson\_number>  
 <start>13:30</start>  
 <end>15:00</end>  
 </time>  
 <room>  
 <id>1419</id>  
 <name>ауд. Lemon Classroom</name>  
 </room>  
 <campus>  
 <id>2</id>  
 <name>Кронверкский пр., д.49, лит.А</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>0</type\_id>  
 <type\_name>Лекция</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 <timetable\_element>  
 <title>История российской науки и техники</title>  
 <group>ИНТ 2.2</group>  
 <teacher>  
 <id>340852</id>  
 <name>Белоусов Александр Сергеевич</name>  
 </teacher>  
 <date>  
 <day>4</day>  
 <weeks>  
 <weeks\_element>1</weeks\_element>  
 <weeks\_element>3</weeks\_element>  
 <weeks\_element>5</weeks\_element>  
 <weeks\_element>7</weeks\_element>  
 <weeks\_element>9</weeks\_element>  
 <weeks\_element>11</weeks\_element>  
 <weeks\_element>13</weeks\_element>  
 <weeks\_element>15</weeks\_element>  
 </weeks>  
 </date>  
 <time>  
 <lesson\_number>5</lesson\_number>  
 <start>15:20</start>  
 <end>16:50</end>  
 </time>  
 <room>  
 <id>2414</id>  
 <name>ауд. 2414</name>  
 </room>  
 <campus>  
 <id>2</id>  
 <name>Кронверкский пр., д.49, лит.А</name>  
 </campus>  
 <lesson\_type>  
 <type\_id>1</type\_id>  
 <type\_name>Практика</type\_name>  
 </lesson\_type>  
 <lesson\_format>  
 <format\_id>1</format\_id>  
 <format\_name>Очно - дистанционный</format\_name>  
 </lesson\_format>  
 </timetable\_element>  
 </timetable>  
</root>

### Сравнение результатов [задания 2](#_Результат_4) и [задания 5](#_Результат_5)

Входной и выходной файл для задания 5 содержат в себе больше данных, чем файлы задания 2. Парсер задания 5 в отличии от парсера задания 2 может обратотать пары ключ-значение, знакение которых равно null или имеет тип bool

## Задание 6

### Код

from time import time  
from os import path  
from main\_task.parser import parse\_json as pj1  
from main\_task.parser.converter import convert\_to\_xml as cx1  
from additional1.main import convert\_to\_xml as cx2  
from additional1.main import parse\_json as pj2  
from additional2.parser import parse\_json as pj3  
from additional2.parser.converter import convert\_to\_xml as cx3  
from additional3.parser import parse\_json as pj4  
from additional3.parser.converter import convert\_to\_xml as cx4  
  
  
def test(converter, parser, string):  
 stime = time()  
 for i in range(100):  
 converter(parser(string))  
 return time() - stime  
  
  
if \_\_name\_\_ == '\_\_main\_\_':  
 json = open(path.join(path.dirname(\_\_file\_\_), "../data/in.json")).read()  
 json1 = open(path.join(path.dirname(\_\_file\_\_), "../data/in-add3.json")).read()  
 print(f"Собственный: {test(cx1, pj1, json)}")  
 print(f"Собственный c регулярными: {test(cx2, pj2, json)}")  
 print(f"С библиотеками: {test(cx3, pj3, json)}")  
 print(f"Собственный полный: {test(cx4, pj4, json1)}")

### Результат

### Объяснение результата

Первый код выполняется быстрее всего, так как он обладает меньшими возможностями, чем коды 3 и 4 и запущен на тесте меньшего размера. Второй код обрабатывается дольше чем первый (хоть у них одинаковые возможности), так как каждое регулярное выражение медленно обрабатывает всю строку на поиск всех “мэтчей”. Третий код и запущен на меньшем объеме данных, поэтому он выполняется быстрее кода 4. Четвертый код выполняется дольше всех, так как обладает самыми большими возможностями (=3) и запущен на самом большом наборе данных.

## Задание 7

Выбранный формат: CSV

### Код

main: <https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/additional5/main.py>

from parser import parse\_json  
from parser.converter import convert\_to\_csv  
from os import path  
  
json = open(path.join(path.dirname(\_\_file\_\_), "../data/in.json"), 'r').read()  
csv = open(path.join(path.dirname(\_\_file\_\_), "../data/out-add5.csv"), 'w')  
csv.write(convert\_to\_csv(parse\_json(json)))

parser: <https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/additional5/parser/__init__.py>

def parse\_json(json: str):  
 string = parse\_value(json.strip())[0]  
 return string  
  
  
def parse\_keyvalue(string: str):  
 parsed\_string = parse\_string(string)  
 if parsed\_string is None:  
 return None  
 parsed\_colon = parse\_colon(parsed\_string[1])  
 if parsed\_colon is None:  
 return None  
 parsed\_value = parse\_value(parsed\_colon[1])  
 if parsed\_value is None:  
 return None  
 return (parsed\_string[0], parsed\_value[0]), parsed\_value[1].strip()  
  
  
def parse\_value(string: str):  
 parsed\_number = parse\_number(string)  
 if not parsed\_number is None:  
 parse\_value(parsed\_number[1].strip())  
 return parsed\_number[0], parsed\_number[1].strip()  
 parsed\_string = parse\_string(string)  
 if not parsed\_string is None:  
 parse\_value(parsed\_string[1].strip())  
 return parsed\_string[0], parsed\_string[1].strip()  
 parsed\_array = parse\_array(string)  
 if not parsed\_array is None:  
 parse\_value(parsed\_array[1].strip())  
 return parsed\_array[0], parsed\_array[1].strip()  
 parsed\_obect = parse\_object(string)  
 if not parsed\_obect is None:  
 parse\_value(parsed\_obect[1].strip())  
 return parsed\_obect[0], parsed\_obect[1].strip()  
  
  
def parse\_comma(string: str):  
 if not string.startswith(','):  
 return None  
 return ",", string[1:].strip()  
  
  
def parse\_comma\_values(string: str):  
 res = []  
 while True:  
 parsed\_value = parse\_value(string)  
 if parsed\_value is None:  
 break  
 res.append(parsed\_value[0])  
 string = parsed\_value[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1]  
 return res, string.strip()  
  
  
def parse\_array(string: str):  
 if not string.startswith('['):  
 return None  
 parsed\_comma\_values = parse\_comma\_values(string[1:].strip())  
 if parsed\_comma\_values is None:  
 res, string = [], string[1:]  
 else:  
 res, string = parsed\_comma\_values  
 if not string.startswith(']'):  
 return None  
 return res, string[1:].strip()  
  
  
def parse\_string(string: str):  
 if not string.startswith("\""):  
 return None  
 second\_quote = string.find("\"", 1)  
 if second\_quote == -1:  
 return None  
 return string[1:second\_quote], string[second\_quote + 1:]  
  
  
def parse\_colon(string: str):  
 if not string.startswith(":"):  
 return None  
 else:  
 return ":", string[2:].strip()  
  
  
def parse\_comma\_keyvalues(string: str):  
 res = {}  
 while True:  
 parsed\_keyvalue = parse\_keyvalue(string)  
 if parsed\_keyvalue is None:  
 break  
 key, val = parsed\_keyvalue[0]  
 res[key] = val  
 string = parsed\_keyvalue[1]  
  
 parsed\_comma = parse\_comma(string)  
 if parsed\_comma is None:  
 break  
 string = parsed\_comma[1].strip()  
 return res, string.strip()  
  
  
def parse\_object(string: str):  
 if not string.startswith("{"):  
 return None  
 arr, string = parse\_comma\_keyvalues(string[1:].strip())  
 if not string.startswith("}"):  
 return None  
 return arr, string[1:].strip()  
  
  
def parse\_number(string: str):  
 num = ""  
 for i in string:  
 if i in "0123456789.Ee-+":  
 num += i  
 else:  
 break  
 if num.isnumeric():  
 return int(num), string[len(num):].strip()  
 else:  
 try:  
 return float(num), string[len(num):].strip()  
 except ValueError:  
 return None

converter: <https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/additional5/parser/converter/__init__.py>

def convert\_to\_csv(d: dict, tag="root"):  
 csv = "title;group;teacher/name;teacher/id;date/day;date/weeks/0;date/weeks/1;date/weeks/2;date/weeks/3;date/weeks/4;date/weeks/5;date/weeks/6;date/weeks/7;date/weeks/8;date/weeks/9;date/weeks/10;date/weeks/11;date/weeks/12;date/weeks/13;date/weeks/14;date/weeks/15;time/lesson\_number;time/start;time/end;room/name;room/id;campus/name;campus/id;lesson\_type/type\_name;lesson\_type/type\_id;lesson\_format/format\_name;lesson\_format/format\_id\n"  
  
 for string in d["timetable"]:  
 res = []  
 res.append(string["title"])  
 res.append(string["group"])  
 res.append(string["teacher"]["name"])  
 res.append(string["teacher"]["id"])  
 res.append(string["date"]["day"])  
  
 for i in string["date"]["weeks"]:  
 res.append(str(i))  
 for i in range(8 - len(string["date"]["weeks"])):  
 res.append(";")  
  
 res.append(string["time"]["lesson\_number"])  
 res.append(string["time"]["start"])  
 res.append(string["time"]["end"])  
 res.append(string["room"]["name"])  
 res.append(string["room"]["id"])  
 res.append(string["campus"]["name"])  
 res.append(string["campus"]["id"])  
 res.append(string["lesson\_type"]["type\_name"])  
 res.append(string["lesson\_type"]["type\_id"])  
 res.append(string["lesson\_format"]["format\_name"])  
 res.append(string["lesson\_format"]["format\_id"])  
  
 csv += ";".join(list(map(str, res))) + "\n"  
  
 return csv

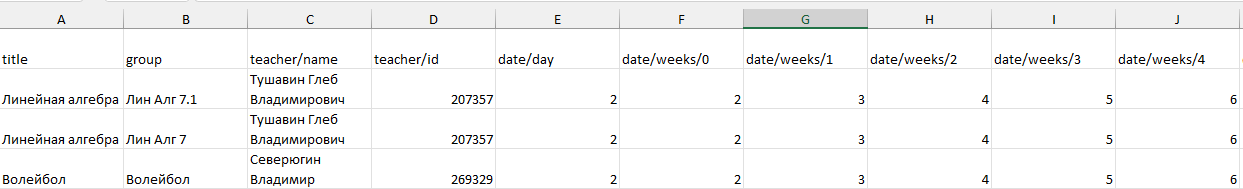
### Результат

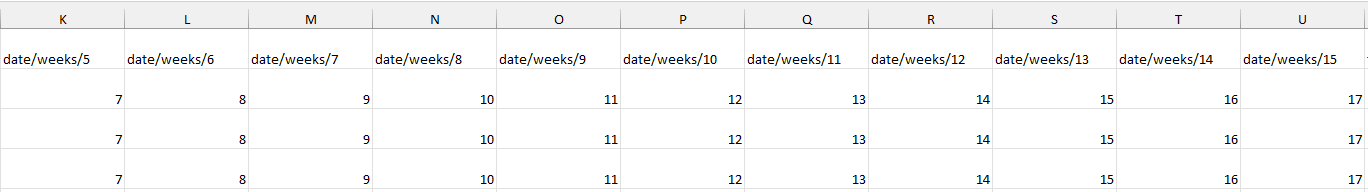
<https://github.com/ldpst/itmo_secret/blob/main/labs/sem1/inf/lab4/tasks/data/out-add5.csv>

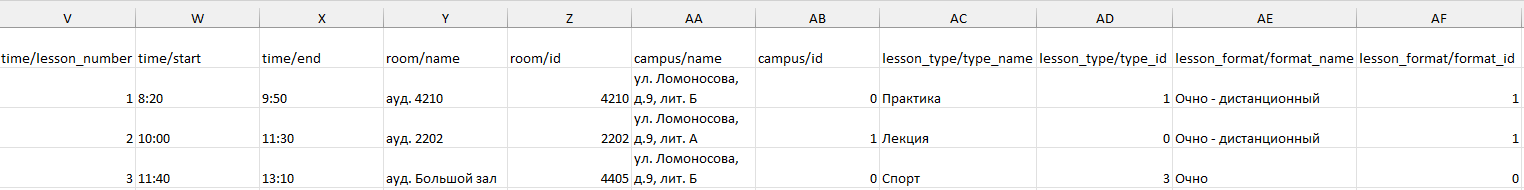
title;group;teacher/name;teacher/id;date/day;date/weeks/0;date/weeks/1;date/weeks/2;date/weeks/3;date/weeks/4;date/weeks/5;date/weeks/6;date/weeks/7;date/weeks/8;date/weeks/9;date/weeks/10;date/weeks/11;date/weeks/12;date/weeks/13;date/weeks/14;date/weeks/15;time/lesson\_number;time/start;time/end;room/name;room/id;campus/name;campus/id;lesson\_type/type\_name;lesson\_type/type\_id;lesson\_format/format\_name;lesson\_format/format\_id  
Линейная алгебра;Лин Алг 7.1;Тушавин Глеб Владимирович;207357;2;2;3;4;5;6;7;8;9;10;11;12;13;14;15;16;17;1;8:20;9:50;ауд. 4210;4210;ул. Ломоносова, д.9, лит. Б;0;Практика;1;Очно - дистанционный;1  
Линейная алгебра;Лин Алг 7;Тушавин Глеб Владимирович;207357;2;2;3;4;5;6;7;8;9;10;11;12;13;14;15;16;17;2;10:00;11:30;ауд. 2202;2202;ул. Ломоносова, д.9, лит. A;1;Лекция;0;Очно - дистанционный;1  
Волейбол;Волейбол;Северюгин Владимир Олегович;269329;2;2;3;4;5;6;7;8;9;10;11;12;13;14;15;16;17;3;11:40;13:10;ауд. Большой зал;4405;ул. Ломоносова, д.9, лит. Б;0;Спорт;3;Очно;0

В виде таблицы:

Таблица разделена на несколько изображений в связи с размером: [рис. 1](#Закладка5), [рис. 2](#Закладка6) и [рис. 3](#Закладка7)

*Рисунок 1 - Таблица. Часть 1*

*Рисунок 2 - Таблица. Часть 2*

*Рисунок 3 - Таблица. Часть 3*

### Описание формата

CSV - табличный формат, разделителем которого служит символ “;”

# Заключение

В процессе выполнения лабораторной работы я познакомился с такими форматами, как JSON, XML, YAML, CSV, TSV. Я написал собственный парсер из JSON файла и конвертер в XML и CSV. Я изучил библиотеки, с помощью которых можно с легкостью обрабатывать данные файлы и углубился в познании регулярных выражений, используя их на практике. Также я получил опыт работы с хронометражом и анализа полученных данных.

# Список источников

1. Лямин А.В., Череповская Е.Н. Объектно-ориентированное программирование. Компьютерный практикум. – СПб: Университет ИТМО, 2017. – 143 с. – Режим доступа: <https://books.ifmo.ru/file/pdf/2256.pdf>.
2. Пишем изящный парсер на Питоне [Электронный ресурс] Хабр : [сайт]. — URL: <https://habr.com/ru/articles/309242> (дата обращения: 28.10.2024).