

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

Факультет программной инженерии и компьютерной техники

Отчёт о лабораторной работе №4  
Исследование протоколов, форматов обмена информацией и  
языков разметки документов  
Вариант №24

Выполнил:

Дядев Владислав Александрович, гр. Р3131

Проверила:

Авксентьева Е. Ю., к.п.н., доцент

Санкт-Петербург 2024

## Оглавление

Задание.....	3
Основные этапы вычисления .....	7
Обязательное задание .....	7
Дополнительное задание №1 .....	7
Дополнительное задание №2 .....	8
Дополнительное задание №3 .....	8
Дополнительное задание №4 .....	8
Дополнительное задание №5 .....	9
Заключение .....	10
Список использованных источников .....	11

## Задание

1. Определить номер варианта как остаток деления на 36 последних двух цифр своего идентификационного номера в ISU: например, 125598 / 36 = 26. В случае, если в оба указанных дня недели нет занятий, то увеличить номер варианта на восемь. В случае, если занятий нет и в новом наборе дней, то продолжать увеличивать на восемь.

2. Изучить форму Бэкуса-Наура.

3. Изучить основные принципы организации формальных грамматик.

4. Изучить особенности языков разметки/форматов JSON, YAML, XML.

5. Понять устройство страницы с расписанием на примере расписания лектора:

[https://itmo.ru/ru/schedule/3/125598/raspisanie\\_zanyatiy.htm](https://itmo.ru/ru/schedule/3/125598/raspisanie_zanyatiy.htm)

6. 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.

7. 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.

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

9. 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.

10. 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.

11. Дополнительное задание №3 (позволяет набрать +25 процентов от максимального числа баллов БаРС за данную лабораторную).

- a) 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 No1.
- b) Проверку осуществить как минимум для расписания с двумя учебными днями по два занятия в каждом.
- c) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

12. 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.

13. 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.

14. Проверить, что все пункты задания выполнены и выполнены верно.

15. Написать отчёт о проделанной работе.

16. Подготовиться к устным вопросам на защите.

Вариант задания представлен на Рисунок 1.

#### 4.4 Варианты заданий

№ варианта	Исходный формат	Результирующий формат	Дни недели
0	JSON	XML	Понедельник, четверг
1	XML	JSON	Понедельник, четверг
2	JSON	YAML	Понедельник, четверг
3	YAML	JSON	Понедельник, четверг
4	YAML	XML	Понедельник, четверг
5	XML	YAML	Понедельник, четверг
6	JSON	XML	Вторник, пятница
7	XML	JSON	Вторник, пятница
8	JSON	YAML	Вторник, пятница
9	YAML	JSON	Вторник, пятница
10	YAML	XML	Вторник, пятница
11	XML	YAML	Вторник, пятница
12	JSON	XML	Среда, суббота
13	XML	JSON	Среда, суббота
14	JSON	YAML	Среда, суббота
15	YAML	JSON	Среда, суббота
16	YAML	XML	Среда, суббота
17	XML	YAML	Среда, суббота
18	JSON	XML	Вторник, четверг
19	XML	JSON	Вторник, четверг
20	JSON	YAML	Вторник, четверг
21	YAML	JSON	Вторник, четверг
22	YAML	XML	Вторник, четверг
23	XML	YAML	Вторник, четверг
24	JSON	XML	Среда, пятница
25	XML	JSON	Среда, пятница
26	JSON	YAML	Среда, пятница
27	YAML	JSON	Среда, пятница
28	YAML	XML	Среда, пятница
29	XML	YAML	Среда, пятница

Рисунок 1 - Вариант задания

## **Основные этапы вычисления**

### **Обязательное задание**

Расписание, составленное на основе варианта, представлено в формате JSON в репозитории на github:

<https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/in.json>

Программа, осуществляющая парсинг из JSON и конвертацию в XML, представлена в репозитории на github:

<https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/main.py>

Итоговый файл с расписанием в формате XML представлен в репозитории на github:

[https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/out\\_main.xml](https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/out_main.xml)

### **Дополнительное задание №1**

Программа, осуществляющая парсинг из JSON и конвертацию в XML с помощью готовой библиотеки, представлена в репозитории на github:

[https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add\\_task1.py](https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add_task1.py)

Итоговый файл с расписанием в формате XML представлен в репозитории на github:

<https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/out1.xml>

Результаты работы этой программы отличается от результата работы исходной тем, что вместо тега `<root></root>` используется тег `<all></all>`, а также другим тегом обозначаются элементы списков (`<item></item>` вместо `<{name_of_list}_item></{name_of_list}_item>`).

### **Дополнительное задание №2**

Программа, осуществляющая парсинг из JSON и конвертацию в XML с использованием регулярных выражений, представлена в репозитории на github:

[https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add\\_task2.py](https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add_task2.py)

Итоговый файл с расписанием в формате XML представлен в репозитории на github:

<https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/out2.xml>

Результат работы этой программы совпадает с результатом работы исходной.

### **Дополнительное задание №3**

Программа, осуществляющая парсинг из JSON и конвертацию в XML с использованием формальных грамматик, представлена в репозитории на github:

[https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add\\_task3.py](https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add_task3.py)

Итоговый файл с расписанием в формате XML представлен в репозитории на github:

<https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/out3.xml>

Результат работы этой программы совпадает с результатом работы исходной.

### **Дополнительное задание №4**

Программа, осуществляющая сравнения стократного времени выполнения предыдущих программ, представлена в репозитории на github:

[https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add\\_task4.py](https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add_task4.py)

Результат работы этой программы представлен на *Рисунок 2*.



```
Time for main.py - 0.0065691471099853516 seconds
Time for add_task1.py - 0.11201930046081543 seconds
Time for add_task2.py - 0.0013041496276855469 seconds
Time for add_task3.py - 0.0010750293731689453 seconds
```

*Рисунок 2 - Время работы программ*

Можно заметить, что быстрее всего работает программа с использованием формальных грамматик, затем идёт программа с использованием регулярных выражений, а после него исходная программа ручных парсинга и конвертации. У первых двух время примерно одинаковое. У исходной программы время также близко к этим двум, потому что все они выполняют ручные парсинг и конвертацию. Дольше же всего работает программа с использованием сторонней библиотеки, потому что в ней учтены все тонкости и случаи парсинга JSON и конвертации его в XML.

#### **Дополнительное задание №5**

Программа, осуществляющая парсинг из JSON и конвертацию в CSV, представлена в репозитории на github:

[https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add\\_task5.py](https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/add_task5.py)

Итоговый файл с расписанием в формате CSV представлен в репозитории на github:

<https://github.com/Alvas07/ITMO/blob/main/1%20Informatics/Labs/Lab4/data/out5.csv>

CSV – это формат таблицы, где в роли заголовков выступают ключи из словаря формат JSON, а в ячейки же записываются значения соответствующих ключей. Каждая ячейка таблицы в файле отделяется символ «;».

## **Заключение**

В ходе лабораторной работы мы изучили формальные грамматики, научились осуществлять парсинг из JSON и конвертирование в XML, а также сравнивать время работы программ.

### **Список использованных 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. <https://habr.com/ru/post/309242/>