

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

*Факультет программной инженерии и компьютерной техники*

**лабораторная работа №4**

**Вариант №18**

**Исследование протоколов, форматов обмена информацией и  
языков разметки документов.**

Выполнил: Пивоваров Р. Н.

Группа: Р3131

Проверил: Авксентьева Е. Ю.

К.п.н. Доцент

Г. Санкт-Петербург, 2024 г.

## **Оглавление**

<b>Задание .....</b>	<b>3</b>
<b>Основные этапы выполнения .....</b>	<b>5</b>
Основное задание .....	5
Дополнительное задание 1 .....	5
Дополнительное задание 2 .....	5
Дополнительное задание 3 .....	5
Дополнительное задание 4 .....	6
Дополнительное задание 5 .....	6
<b>Заключение.....</b>	<b>7</b>
<b>Список литературы.....</b>	<b>7</b>

**лабораторная работа №4**  
**Исследование протоколов, форматов обмена информацией и**  
**языков разметки документов.**

## 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). а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов. б) Переписать исходный код, применив найденные библиотеки.

Регулярные выражения также нельзя использовать. 1 с) 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). а) Переписать исходный код, добавив в него использование регулярных выражений. б) 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). 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. б) 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.

12. Дополнительное задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную). а) Используя свою исходную программу из обязательного задания и программы из дополнительных заданий, сравнить стократное время выполнения парсинга + конвертации в цикле. б) 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). а) 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. 2 б) Проанализировать

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

### **Основные этапы выполнения**

#### **Основное задание**

[Ссылка на исходный файл расписания в формате JSON на Github](#)

[Ссылка на листинг программы на Github](#)

[Ссылка на выходной файл программы на Github](#)

#### **Дополнительное задание 1**

[Ссылка на листинг программы на Github](#)

[Ссылка на выходной файл программы на Github](#)

Значительных различий в выводе программ нет, отличается лишь формат вывода массивов, но т.к. оба формата допустимы и я мог написать основную программу с аналогичным выводом массивов, то это различие можно не учитывать. Также можно заметить колоссальное уменьшение количества кода необходимого для преобразования файла. Причиной этого является сокрытие кода парсера в файлах библиотеки.

#### **Дополнительное задание 2**

[Ссылка на листинг программы на Github](#)

[Ссылка на выходной файл программы на Github](#)

Различий в выводе исходной и измененной программ нет. Но следует отметить уменьшение количества кода требуемого для преобразования файла.

#### **Дополнительное задание 3**

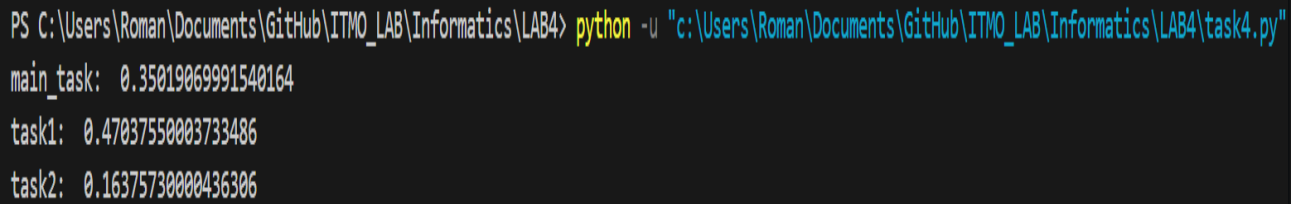
[Ссылка на листинг программы на Github](#)

[Ссылка на выходной файл программы на Github](#)

Различия в работе программы отсутствуют, так как код изначально писался с применением знаний о формальных грамматиках, для полноценного парсинга из JSON в XML.

## Дополнительное задание 4

[Ссылка на листинг программы на Github](#)



```
PS C:\Users\Roman\Documents\Github\ITMO_LAB\Informatics\LAB4> python -u "c:\Users\Roman\Documents\Github\ITMO_LAB\Informatics\LAB4\task4.py"
main_task: 0.35019069991540164
task1: 0.47037550003733486
task2: 0.16375730000436306
```

Рисунок 1 вывод дополнительного задания 4

По Рисунок 1 вывод дополнительного задания 4 мы можем заметить что вариант из дополнительного задания 1 выполняется дольше всего. Я думаю что это связано с тем что их алгоритм реализует больше функций и с тем что на обращение к библиотеке тоже уходит время. К тому же можно заметить что алгоритм с использованием регулярных выражений справляется быстрее. Это связано с тем что мы меньше делаем меньшее количество рекурсий за счет использования регулярных выражений.

## Дополнительное задание 5

[Ссылка на листинг программы на Github](#)

[Ссылка на выходной файл программы на Github](#)

Для данного задания мной был выбран формат WML.

WML (Wireless Markup Language) — это язык разметки, разработанный специально для мобильных устройств с использованием WAP (Wireless Application Protocol). Он был популярен в конце 1990-х и начале 2000-х годов, до широкого распространения современных мобильных браузеров с поддержкой HTML.

В части парсинга из Json в структуры данных Python программа не получила значительных изменений. Изменения появились лишь в части конвертации по понятным причинам. Так теперь ключи dict записываются не как `<key></key>`, а как `<card id="key" title="key"></card>`. Элементы массив же записываются в тег `<p></p>` вместо `<item></item>`.

## **Заключение**

В результате выполнения лабораторной работы я узнал формальные грамматики JSON, XML, WML, научился работать с файлами этих форматов и преобразовывать их из одного формата в другой.

### **Список литературы**

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 / Рецензент: Поляков В.И. - Санкт-Петербург: Университет ИТМО, 2019. - 56 с. - экз. –
2. Грошев А.С. Г89 Информатика: Учебник для вузов / А.С. Грошев. – Архангельск, Арханг. гос. техн. ун-т, 2010. - 470с.