Информатика

Лабораторная работа №4

Вариант №2

**Выполнил:**

Воронов Григорий Алексеевич

Группа Р3116

**Преподаватель:**

Авксентьева Елена Юрьевна

кандидат педагогических наук, доцент факультета ПИиКТ

# Задание

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

**2.** Изучить форму Бэкуса-Наура.

**3.** Изучить особенности языков разметки/форматов JSON, YAML, XML.

**4.** Понять устройство страницы с расписанием на примере расписания лектора:

https://itmo.ru/ru/schedule/3/125598/raspisanie\_zanyatiy.htm

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

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

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

**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 +10 процентов от максимального числа баллов БаРС за данную лабораторную).

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

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

**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).

**c)** 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.

**d)** Проанализировать полученные результаты, объяснить особенности использования формата.

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

**15.** Написать отчёт о проделанной работе.

**16.** Подготовиться к устным вопросам на защите

|  |  |  |  |
| --- | --- | --- | --- |
| 2 | JSON | YAML | Понедельник |

# Основные этапы выполнения

## 1. Обязательное задание

Исходный файл json: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/data/input.json

Исходный код: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/main\_parser.py

Результат: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/data/output.yml

## 2. Дополнительное задание №1

Исходный код: <https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/lib_parser.py>

Результат этого дополнительного задания не отличается от результата обязательного задания, кроме отсутствия некоторых незначительных символов переноса строки, которые не влияют на синтаксис. Оба YAML файла валидны.

## 3. Дополнительное задание №2

Исходный код: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/re\_parser.py

Файл результата не отличается от результата обязательного задания. Единственное изменение в коде программы - парсинг чисел и строк заменены на регулярные выражения.

## 4. Дополнительное задание №3

Исходный код: [https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/main\_parser.py](https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/re_parser.py)

Мной изначально был написан код с использованием формальных грамматик. Поэтому реализация этого дополнительного задания, ровно как и результат выполнения, идентична основному.

## 5. Дополнительное задание №4

Исходный код: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/testing4.py

1) Время работы программы для обязательного задания:

0.1045525074005127 секунды;

2) Время работы программы для доп. задания №1 (программа использует библиотеку yaml, json):

0.12968897819519043 секунды;

3) Время работы программы для доп. задания №2 (программа использует библиотеку re):

0.1005091667175293 секунды.

Время работы (1) и (3) практически одинаковы. Это связано с тем, что их реализация практически идентична, за исключением парсинга чисел и строк (где в доп. задании №2 используются регулярные выражения). Скорость работы парсинга, реализованного при помощи библиотеки, ниже, виною чему, вероятнее всего, являются дополнительные механизмы парсинга и перевода, которые реализованы в самой библиотеке, но избыточны в данном задании.

## 6. Дополнительное задание №5

Исходный код: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/src/to\_csv.py

Результат: https://github.com/grigory222/JustMyStrangeLabs/blob/main/infa/lab4/data/output.csv

**CSV** (*Comma-Separated Values* — значения, разделённые запятыми) — текстовый формат, предназначенный для представления табличных данных. Строка таблицы соответствует строке текста, которая содержит одно или несколько полей, разделенных запятыми.

# Вывод

Во время выполнения лабораторной работы я узнал о языках разметки JSON и YAML, научился с ними работать и переводить один в другой с помощью средств языка Python. Научился работать с некоторыми библиотеками для парсинга, такими как json и yaml; узнал о CSV и научился с ним работать.

# 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. - Режим доступа:

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

1. Грошев А.С. Г89 Информатика: Учебник для вузов / А.С. Грошев. – Архангельск, Арханг. гос. техн. ун-т, 2010. -470с. -Режим доступа <https://narfu.ru/university/library/books/0690.pdf>