

Национальный Исследовательский Университет ИТМО  
Факультет Программной Инженерии и Компьютерной Техники



Лабораторная работа №4 по информатике  
«Исследование языков разметки документов»  
Вариант №7

Работу выполнил:  
Егошин Алексей Васильевич,  
группа Р3113

Преподаватель:  
Малышева Татьяна Алексеевна

Санкт-Петербург, 2019

### ***Краткая формулировка задачи:***

1. Определить номер варианта как остаток деления на 35 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, PROTOBUF.
4. Понять устройство страницы с расписанием для своей группы:  
<http://www.ifmo.ru/ru/schedule/0/P3200/schedule.htm>
5. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного.
6. Написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.
7. Нельзя использовать готовые библиотеки, кроме re (регулярные выражения в Python) и библиотеки для загрузки XML-файлов.
8. Необязательное задание для получения оценки «4» и «5» (позволяет набрать от 75 до 89 процентов от максимального числа баллов БаРС за данную лабораторную).
  - а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
  - б) Сравнить полученные результаты и объяснить их сходство/различие.
9. Необязательное задание для получения оценки «5» (позволяет набрать от 90 до 100 процентов от максимального числа баллов БаРС за данную лабораторную).
  - а) Используя свою программу и найденные готовые библиотеки, сравнить десятикратное время выполнения парсинга + конвертации в цикле.
  - б) Проанализировать полученные результаты и объяснить их сходство/различие.
10. Проверить, что все пункты задания выполнены и выполнены верно.
11. Написать отчёт о проделанной работе.
12. Подготовиться к устным вопросам на защите.

### ***Исходный код:***

<https://github.com/Roggired/info-lab4>

**Входной файл** *example.json*:

```
{
  "day": "Monday",
  "NumberOfLessons": 8,
  "Lesson1": {
    "Time": "10.00 - 11.30",
    "Parity": "YES",
    "Address": "371 Кронверкский пр. д.49 лит.А",
    "Subject": "ПРОГРАММИРОВАНИЕ (Лаб)",
    "Teacher": "Письмак Алексей Евгеньевич"
  },
  "Lesson2": {
    "Time": "10.00 - 11.30",
    "Parity": "NO",
    "Address": "372 Кронверкский пр. д.49 лит.А",
    "Subject": "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ (Лаб)",
    "Teacher": "Перминов Илья Валентинович"
  },
  "Lesson3": {
    "Time": "11.40 - 13.10",
    "Parity": "NO",
    "Address": "371 Кронверкский пр. д.49 лит.А",
    "Subject": "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ (Лаб)",
    "Teacher": "Перминов Илья Валентинович"
  },
  "Lesson4": {
    "Time": "11.40 - 13.10",
    "Parity": "YES",
    "Address": "371 Кронверкский пр. д.49 лит.А",
    "Subject": "ПРОГРАММИРОВАНИЕ (Лаб)",
    "Teacher": "Письмак Алексей Евгеньевич"
  },
  "Lesson5": {
    "Time": "13.30 - 15.00",
    "Parity": "YES",
    "Address": "369А Кронверкский пр. д.49 лит.А",
    "Subject": "ДИСКРЕТНАЯ МАТЕМАТИКА (Прак)",
    "Teacher": "Поляков Владимир Иванович"
  },
  "Lesson6": {
    "Time": "13.30 - 15.00",
    "Parity": "NO",
    "Address": "304 Кронверкский пр. д.49 лит.А",
    "Subject": "ИНФОРМАТИКА (Лаб)",
    "Teacher": "Малышева Татьяна Алексеевна"
  },
  "Lesson7": {
    "Time": "15.20 - 16.50",
    "Parity": "NO",
    "Address": "304 Кронверкский пр. д.49 лит.А",
    "Subject": "ИНФОРМАТИКА (Лаб)",
    "Teacher": "Малышева Татьяна Алексеевна"
  },
  "Lesson8": {
    "Time": "15.20 - 16.50",
    "Parity": "YES",
    "Address": "369А Кронверкский пр. д.49 лит.А",
    "Subject": "ДИСКРЕТНАЯ МАТЕМАТИКА (Прак)",
    "Teacher": "Поляков Владимир Иванович"
  }
}
```

***Выходной файл example.yaml:***

day: "Monday"

NumberOfLessons: 8

Lesson1:

Time: "10.00-11.30"

Parity: "YES"

Address: "371Кронверкскийпр.д.49лит.А"

Subject: "ПРОГРАММИРОВАНИЕ(Лаб)"

Teacher: "ПисьмакАлексейЕвгеньевич"

Lesson2:

Time: "10.00-11.30"

Parity: "NO"

Address: "372Кронверкскийпр.д.49лит.А"

Subject: "ОСНОВЫПРОФЕССИОНАЛЬНОЙДЕЯТЕЛЬНОСТИ(Лаб)"

Teacher: "ПерминовИльяВалентинович"

Lesson3:

Time: "11.40-13.10"

Parity: "NO"

Address: "371Кронверкскийпр.д.49лит.А"

Subject: "ОСНОВЫПРОФЕССИОНАЛЬНОЙДЕЯТЕЛЬНОСТИ(Лаб)"

Teacher: "ПерминовИльяВалентинович"

Lesson4:

Time: "11.40-13.10"

Parity: "YES"

Address: "371Кронверкскийпр.д.49лит.А"

Subject: "ПРОГРАММИРОВАНИЕ(Лаб)"

Teacher: "ПисьмакАлексейЕвгеньевич"

Lesson5:

Time: "13.30-15.00"

Parity: "YES"

Address: "369АКронверкскийпр.д.49лит.А"

Subject: "ДИСКРЕТНАЯМАТЕМАТИКА(Прак)"

Teacher: "ПоляковВладимирИванович"

Lesson6:

Time: "13.30-15.00"

Parity: "NO"

Address: "304Кронверкскийпр.д.49лит.А"

Subject: "ИНФОРМАТИКА(Лаб)"

Teacher: "МальшеваТатьянаАлексеевна"

Lesson7:

Time: "15.20-16.50"

Parity: "NO"

Address: "304Кронверкскийпр.д.49лит.А"

Subject: "ИНФОРМАТИКА(Лаб)"

Teacher: "МальшеваТатьянаАлексеевна"

Lesson8:

Time: "15.20-16.50"

Parity: "YES"

Address: "369АКронверкскийпр.д.49лит.А"

Subject: "ДИСКРЕТНАЯМАТЕМАТИКА(Прак)"

Teacher: "ПоляковВладимирИванович"

### ***Сравнение производительности:***

Проведен эксперимент по измерению десятикратного времени выполнения программы при помощи библиотеки *time*:

Mode 0 – библиотечный парсинг и конвертация.

Mode 1 – собственный парсинг и конвертация.

Измерение 1:

Mode 0 - 0.235 с

Mode 1 - 0.160 с

Измерение 2:

Mode 0 - 0.129 с

Mode 1 - 0.165 с

Измерение 3:

Mode 0 - 0.107 с

Mode 1 - 0.168 с

Измерение 4:

Mode 0 - 0.150 с

Mode 1 - 0.140 с

Измерение 5:

Mode 0 - 0.112 с

Mode 1 - 0.124 с

Среднее время:

Mode 0 – 0,147 с

Mode 1 – 0,151 с

### ***Вывод:***

В ходе лабораторной работы я повторил языки разметки, а также попробовал написать свой собственный json-парсер. В результате ~~много~~мной был написан парсер со следующей допущениями:

1. Входной файл в целом синтаксически корректен. Реализованы лишь базовые проверки синтаксиса на основе регулярных выражений.
2. Отсутствуют вложенные объекты и массивы.

Общий алгоритм работы парсера, очевидно, не является оптимальным в силу вышеизложенных допущений, однако с учетом имеющихся знаний в структурах данных и алгоритмах я считаю, что сложность и компетентность парсера адекватна. ~~(На жабе было бы лучше, жаба лучше змеи!)~~ Результаты сравнения времени работы написанных мной парсера и конвертера с существующими аналогами показывают, что на входных файлах, удовлетворяющих вышеизложенным допущениям, написанный мной код работает в целом оптимально.