

Университет ИТМО
Факультет ПИиКТ

Лабораторная работа №3 по информатике

Исследование языков разметки документов

Вариант 24

Выполнил:

Сантистебан Монтерос Кевин Хансель

группа Р3113

Преподаватель: Балакшин Павел Валерьевич

г. Санкт-Петербург

2020 год

Порядок выполнения работы

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.

https://github.com/Kevin-Santisteban/Informatics_Lab3/blob/master/README.md

Контекстно-свободная грамматика для JSON файлов (Бэкуса-Наура)

<json> ::= <primitive> | <container>

<primitive> ::= <number> | <string> | <boolean>

<container> ::= <object> | <array>

<array> ::= '[' [<json> *(' , ' <json>)] ']'

<object> ::= '{' [<member> *(' , ' <member>)] '}'

<member> ::= <string> ': ' <json>

<number> Допустимое вещественное число, выраженное в одном из нескольких заданных форматов.

<string> Строка допустимых символов, заключенная в кавычки

<boolean> Буквенные строки 'true', 'false' или 'null' (без кавычек)

<array> Последовательность значений, разделенных запятыми

<object> Последовательность <member>

<member> Пара, состоящая из строк и значения <json>

https://github.com/Kevin-Santisteban/Informatics_Lab3/blob/master/README.md

Исходный файл в формате JSON Weekday.json

```
{
  "name": "Пятница",
  "lessons": [
    {
      "name": "Дискретная математика (прак)",
      "place": "Кронверкский пр., д.49, лит.А",
      "classroom": "403",
      "startTime": "10:00",
      "endTime": "11:30",
      "professor": "Поляков Владимир Иванович",
      "presencial": true,
      "online": true,
      "oddWeek": true
    },
    {
      "name": "Дискретная математика (прак)",
      "place": "Кронверкский пр., д.49, лит.А",
      "classroom": "403",
      "startTime": "11:40",
      "endTime": "13:10",
      "professor": "Поляков Владимир Иванович",
      "presencial": true,
      "online": true,
      "oddWeek": true
    },
    {
      "name": "История (лекция)",
      "place": "Кронверкский пр., д.49, лит.А",
      "classroom": "Актальный зал",
      "startTime": "10:00",
      "endTime": "11:30",
      "professor": "Жиркова Галина Петровна",
      "presencial": true,
      "online": true,
      "oddWeek": false
    },
    {
```

```
    "name": "История (прак) ",  
    "place": "Кронверкский пр., д.49, лит.А",  
    "classroom": "461",  
    "startTime": "11:40",  
    "endTime": "13:10",  
    "professor": "Жиркова Галина Петровна",  
    "presencial": true,  
    "online": true,  
    "oddWeek": false  
  }  
]  
}
```

[https://github.com/Kevin-Santisteban/Informatics_Lab3/
blob/master/Weekday.json](https://github.com/Kevin-Santisteban/Informatics_Lab3/blob/master/Weekday.json)

Код программы на питоне Lab3.py

```
def getFileContent(filePath):
    newFile = open(filePath, "r")
    return newFile

def hasNumbers(inputString):
    return any(char.isdigit() for char in inputString)

def parsingJsonFileToYaml():
    # timeFormat = "%H:%M"

    jsonFile = getFileContent("Weekday.json")

    #getting a list of jsonfile lines
    if jsonFile.mode == "r":
        jsonFiletext = jsonFile.readlines()

    #Creating Output file
    outputFile = open("Weekday.yaml", "w")

    lineCount = 0
    for line in jsonFiletext:
        lineCount += 1
        if lineCount == 1:
            line = line.rstrip(" ")
            line = line.replace("{", "---")
            outputFile.write(line)
        elif (lineCount == len(jsonFiletext)):
            line = line.replace("}", "")
            outputFile.write(line)
        else:

            #validating if is not an object
            if ":{ " not in line and ": { " not in line :
```

```

        #Replacing chars, so the syntax matches yaml:
        line=
line.replace("{","-").replace("}\n","").replace("}", "").replace("[","").repl
ace("]", "").replace("},","").replace(",\n","\n")

        tempLine = line.split(": ")
        #Validating that there's something meaningful after colon(:)
        if len(tempLine)>1:
            #Removing quotes from left side :
            line = tempLine[0].replace("\"","") + " : "+
str(tempLine[1])
        else:
            tempLine2 = line.strip().split(":")
            line = tempLine2[0].replace("\"","").replace("{","") + ":\n"

        if len(line.lstrip())!=0:
            outputFile.write(line)

def main():
    parsingJsonFileToYaml()
    return

if __name__ == "__main__":
    main()
    print("End of execution")

```

https://github.com/Kevin-Santisteban/Informatics_Lab3/blob/master/Lab3.py

Результирующий выходной файл в формате YAML Weekday.yaml

```
---
name : "Пятница"
lessons :
  -
    name : "Дискретная математика (прак)"
    place : "Кронверкский пр., д.49, лит.А"
    classroom : "403"
    startTime : "10:00"
    endTime : "11:30"
    professor : "Поляков Владимир Иванович"
    presencial : true
    online : true
    oddWeek : true
  -
    name : "Дискретная математика (прак)"
    place : "Кронверкский пр., д.49, лит.А"
    classroom : "403"
    startTime : "11:40"
    endTime : "13:10"
    professor : "Поляков Владимир Иванович"
    presencial : true
    online : true
    oddWeek : true
  -
    name : "История (лекция)"
    place : "Кронверкский пр., д.49, лит.А"
    classroom : "Актный зал"
    startTime : "10:00"
    endTime : "11:30"
    professor : "Жиркова Галина Петровна"
    presencial : true
    online : true
    oddWeek : false
  -
    name : "История (прак)"
```



```
place : "Кронверкский пр., д.49, лит.А"  
classroom : "461"  
startTime : "11:40"  
endTime : "13:10"  
professor : "Жиркова Галина Петровна"  
presencial : true  
online : true  
oddWeek : false
```

[https://github.com/Kevin-Santisteban/Informatics_Lab3/
blob/master/Weekday.yaml](https://github.com/Kevin-Santisteban/Informatics_Lab3/blob/master/Weekday.yaml)

Код программы на питоне с библиотеками Lab3Libs.py

```
import time

def getFileContent(filePath):
    newFile = open(filePath, "r")
    return newFile

def hasNumbers(inputString):
    return any(char.isdigit() for char in inputString)

def parsingJsonFileToYaml():
    # timeFormat = "%H:%M"

    jsonFile = getFileContent("Weekday.json")

    #getting a list of jsonfile lines
    if jsonFile.mode == "r":
        jsonFiletext = jsonFile.readlines()

    #Creating Output file
    outputFile = open("Weekday.yaml", "w")

    lineCount = 0
    for line in jsonFiletext:
        lineCount += 1
        if lineCount == 1:
            line = line.rstrip(" ")
            line = line.replace("{", "---")
            outputFile.write(line)
        elif(lineCount == len(jsonFiletext)):
            line = line.replace("}", "")
            outputFile.write(line)
        else:
```

```

        #validating if is not an object
        if ":{ " not in line and ": {" not in line :

            #Replacing chars, so the syntax matches yaml:
            line=
line.replace("{","-").replace("}\n","").replace("}","").replace("[
","").replace("]", "").replace("}","").replace(",\n","\n")

            tempLine = line.split(": ")
            #Validating that there's something meaningful after
colon(:)

            if len(tempLine)>1:
                #Removing quotes from left side :
                line = tempLine[0].replace("\"","") + " : "+
str(tempLine[1])
            else:
                tempLine2 = line.strip().split(":")
                line =
tempLine2[0].replace("\"","").replace("{","") + ":\n"

            if len(line.lstrip())!=0:
                outputFile.write(line)

def main():
    parsingJsonFileToYaml()
    return

if __name__ == "__main__":
    start_time = time.time()
    main()
    print("---No Libs %s seconds ---" % (time.time() - start_time))

```

https://github.com/Kevin-Santisteban/Informatics_Lab3/blob/master/Lab3Libs.py

Результирующий выходной файл в формате YAML LibWeekday.yaml

```
lessons:
- endTime: '11:30'
  name: Дискретная математика (прак)
  oddWeek: true
  online: true
  place: Кронверкский пр., д.49, лит.А
  presencial: true
  professor: Поляков Владимир Иванович
  startTime: '10:00'
  classroom: '403'
- endTime: '13:10'
  name: Дискретная математика (прак)
  oddWeek: true
  online: true
  place: Кронверкский пр., д.49, лит.А
  presencial: true
  professor: Поляков Владимир Иванович
  startTime: '11:40'
  classroom: '403'
- endTime: '11:30'
  name: История (лекция)
  oddWeek: false
  online: true
  place: Кронверкский пр., д.49, лит.А
  presencial: true
  professor: Жиркова Галина Петровна
  startTime: '10:00'
  classroom: Актовый зал
- endTime: '13:10'
  name: История (прак)
  oddWeek: false
  online: true
  place: Кронверкский пр., д.49, лит.А
  presencial: true
  professor: Жиркова Галина Петровна
```

```
startTime: '11:40'  
classroom: '461'  
name: Пятница
```

Сравнение времени выполнения

Lab3.py

---No Libs 0.0020532608032226562 seconds

Lab3Libs.py

---With Libs 0.004752397537231445 seconds ---

Вывод

В выполнении лабораторной работы можно было узнать о различных форматах файлов, и создать прототип парсера JSON файла, переводящий его в формат YAML.

Созданный парсер занял меньше времени чем готовые библиотеки