

Национальный исследовательский университет информационных технологий,
механики и оптики

Факультет программной инженерии и компьютерной техники

Информатика
Лабораторная работа №3
«Исследование языков разметки документов»

Работу выполнил:

Бавыкин Роман

Преподаватель:

Балакшин Павел Валерьевич

Группа: Р3110

Вариант 4

Санкт-Петербург

2020 г.

Задача:

1. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием для своей группы.
2. Написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.

Необязательные задания:

1. а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
б) Сравнить полученные результаты и объяснить их.
2. а) Используя свою программу и найденные готовые библиотеки, сравнить десятикратное время выполнения парсинга + конвертации в цикле.
б) Проанализировать полученные результаты и объяснить их.

Листинг программы .py:

```
1 inputFile = open('schedule.json', 'r')
2 outputFile = open('schedule.xml', 'w')
3 lines = list()
4 counter = -1
5 a = ['', '', '', '', '', '']
6 nextLine = inputFile.readline()
7 nextLine = inputFile.readline()
8 while nextLine:
9     lines.append(nextLine)
10    nextLine = inputFile.readline()
11 print(lines)
12 outputFile.write('<?xml version="1.0" ?>\n<all>\n')
13 for line in lines:
14     if (line.find('{') > -1):
15         line = line[:line.find('{')] + '<' + line[line.find('{') + 1:line.rfind('}')] + ' type="dict">' + '\n'
16         outputFile.write(line)
17         counter += 1
18         a[counter] = line
19     elif (line.find('" : "') > -1):
20         line = line[:line.find('" : "') + 1:]
21         line = line[:line.find('" : "') + 1:] + ' type="str">' + line[line.find('" : "') + 1:]
22         line = line[:line.find('" : "') + 1:] + line[line.find('" : "') + 1:]
23         line = line[:line.find('" : "') + 1:] + '</' + line[line.find('" : "') + 1:line.rfind(' type="str">')] + '>\n'
24         line = line.replace('str', "str")
25         outputFile.write(line)
26     elif (line == ' '):
27         outputFile.write('</all>')
28     elif (line.find('}') > -1):
29         line = line[:line.find('}')] + '</' + a[counter][a[counter].find('<')]
30         + 1:a[counter].find(' type="dict">')] + '>\n'
31         counter -= 1
32         outputFile.write(line)
```

Содержание файла в исходном формате (JSON):

```
1 {
2     "Понедельник" : {
3         "НечетнаяНеделя" : {
4             "Пара1-1" : {
5                 "Время" : "08:20-09:50",
6                 "Аудитория" : "303/2 (усл)",
7                 "Корпус" : "Кронверкский пр., д.49, лит.А",
8                 "Предмет" : "ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ",
9                 "Преподаватель" : "Сорокин Роман Борисович"
10            },
11            "Пара1-2" : {
12                "Время" : "08:20-09:50",
13                "Аудитория" : "305",
14                "Корпус" : "Кронверкский пр., д.49, лит.А",
15                "Предмет" : "ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ",
16                "Преподаватель" : "Горбунов Михаил Витальевич"
17            },
18            "Пара2-1" : {
19                "Время" : "10:00-11:30",
20                "Аудитория" : "303/2 (усл)",
21                "Корпус" : "Кронверкский пр., д.49, лит.А",
22                "Предмет" : "ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ",
23                "Преподаватель" : "Сорокин Роман Борисович"
24            },
25            "Пара2-2" : {
26                "Время" : "10:00-11:30",
27                "Аудитория" : "305",
28                "Корпус" : "Кронверкский пр., д.49, лит.А",
29                "Предмет" : "ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ",
30                "Преподаватель" : "Горбунов Михаил Витальевич"
31            },
32            "Пара3" : {
33                "Время" : "11:40-13:10",
34                "Аудитория" : "369А",
35                "Корпус" : "Кронверкский пр., д.49, лит.А",
36                "Предмет" : "ДИСКРЕТНАЯ МАТЕМАТИКА (ПРАК) ",
37                "Преподаватель" : "Поляков Владимир Иванович"
38            },
39            "Пара4" : {
40                "Время" : "13:30-15:00",
41                "Аудитория" : "369А",
42                "Корпус" : "Кронверкский пр., д.49, лит.А",
43                "Предмет" : "ДИСКРЕТНАЯ МАТЕМАТИКА (ПРАК) ",
44                "Преподаватель" : "Поляков Владимир Иванович"
45            }
46        },
47        "ЧетнаяНеделя" : {
48            "Пара1" : {
49                "Время" : "08:20-09:50",
50                "Аудитория" : "303/2 (усл)",
51                "Корпус" : "Кронверкский пр., д.49, лит.А",
52                "Предмет" : "ИНФОРМАТИКА (ЛАБ) ",
53                "Преподаватель" : "Балакшин Павел Валерьевич"
54            },
55            "Пара2" : {
56                "Время" : "10:00-11:30",
57                "Аудитория" : "303/2 (усл)",
58                "Корпус" : "Кронверкский пр., д.49, лит.А",
59                "Предмет" : "ИНФОРМАТИКА (ЛАБ) ",
60                "Преподаватель" : "Балакшин Павел Валерьевич"
61            },
62        }
63    }
```

```

62 "Пара3-1" : {
63     "Время" : "11:40-13:10",
64     "Аудитория" : "305",
65     "Корпус" : "Кронверкский пр., д.49, лит.А",
66     "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ) : 1/2 ГРУППЫ",
67     "Преподаватель" : "Яркеев Александр Сергеевич"
68 },
69 "Пара3-2" : {
70     "Время" : "11:40-13:10",
71     "Аудитория" : "303/2 (УСЛ)",
72     "Корпус" : "Кронверкский пр., д.49, лит.А",
73     "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ) : 1/2 ГРУППЫ",
74     "Преподаватель" : "Ткешелашвили Нино Мирабиевна"
75 },
76 "Пара4-1" : {
77     "Время" : "13:30-15:00",
78     "Аудитория" : "305",
79     "Корпус" : "Кронверкский пр., д.49, лит.А",
80     "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ) : 1/2 ГРУППЫ",
81     "Преподаватель" : "Яркеев Александр Сергеевич"
82 },
83 "Пара4-2" : {
84     "Время" : "13:30-15:00",
85     "Аудитория" : "303/2 (УСЛ)",
86     "Корпус" : "Кронверкский пр., д.49, лит.А",
87     "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ) : 1/2 ГРУППЫ",
88     "Преподаватель" : "Ткешелашвили Нино Мирабиевна"
89 }
90 }
91 }
92 }

```

Содержание файла в результирующем формате (XML):

```

1  <?xml version="1.0" ?>
2  <all>
3      <Понедельник type="dict">
4          <НечетнаяНеделя type="dict">
5              <Пара1-1 type="dict">
6                  <Время type="str">08:20-09:50</Время>
7                  <Аудитория type="str">303/2 (усл)</Аудитория>
8                  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
9                  <Предмет type="str">ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ</Предмет>
10                 <Преподаватель type="str">Сорокин Роман Борисович</Преподаватель>
11             </Пара1-1>
12             <Пара1-2 type="dict">
13                 <Время type="str">08:20-09:50</Время>
14                 <Аудитория type="str">305</Аудитория>
15                 <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
16                 <Предмет type="str">ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ</Предмет>
17                 <Преподаватель type="str">Горбунов Михаил Витальевич</Преподаватель>
18             </Пара1-2>
19             <Пара2-1 type="dict">
20                 <Время type="str">10:00-11:30</Время>
21                 <Аудитория type="str">303/2 (усл)</Аудитория>
22                 <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
23                 <Предмет type="str">ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ</Предмет>
24                 <Преподаватель type="str">Сорокин Роман Борисович</Преподаватель>
25             </Пара2-1>
26             <Пара2-2 type="dict">
27                 <Время type="str">10:00-11:30</Время>
28                 <Аудитория type="str">305</Аудитория>
29                 <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
30                 <Предмет type="str">ПРОГРАММИРОВАНИЕ (ЛАБ) : 1/2 ГРУППЫ</Предмет>

```

```

31         <Преподаватель type="str">Горбунов Михаил Витальевич</Преподаватель>
32     </Пара2-2>
33     <Пара3 type="dict">
34         <Время type="str">11:40-13:10</Время>
35         <Аудитория type="str">369А</Аудитория>
36         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
37         <Предмет type="str">ДИСКРЕТНАЯ МАТЕМАТИКА (ПРАК)</Предмет>
38         <Преподаватель type="str">Поляков Владимир Иванович</Преподаватель>
39     </Пара3>
40     <Пара4 type="dict">
41         <Время type="str">13:30-15:00</Время>
42         <Аудитория type="str">369А</Аудитория>
43         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
44         <Предмет type="str">ДИСКРЕТНАЯ МАТЕМАТИКА (ПРАК)</Предмет>
45         <Преподаватель type="str">Поляков Владимир Иванович</Преподаватель>
46     </Пара4>
47 </НечетнаяНеделя>
48 <ЧетнаяНеделя type="dict">
49     <Пара1 type="dict">
50         <Время type="str">08:20-09:50</Время>
51         <Аудитория type="str">303/2 (усл)</Аудитория>
52         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
53         <Предмет type="str">ИНФОРМАТИКА (ЛАБ)</Предмет>
54         <Преподаватель type="str">Балакшин Павел Валерьевич</Преподаватель>
55     </Пара1>
56     <Пара2 type="dict">
57         <Время type="str">10:00-11:30</Время>
58         <Аудитория type="str">303/2 (усл)</Аудитория>
59         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
60         <Предмет type="str">ИНФОРМАТИКА (ЛАБ)</Предмет>
61         <Преподаватель type="str">Балакшин Павел Валерьевич</Преподаватель>
62     </Пара2>
63     <Пара3-1 type="dict">
64         <Время type="str">11:40-13:10</Время>
65         <Аудитория type="str">305</Аудитория>
66         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
67         <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>
68         <Преподаватель type="str">Яркеев Александр Сергеевич</Преподаватель>
69     </Пара3-1>
70     <Пара3-2 type="dict">
71         <Время type="str">11:40-13:10</Время>
72         <Аудитория type="str">303/2 (УСЛ)</Аудитория>
73         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
74         <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>
75         <Преподаватель type="str">Ткешелашвили Нино Мирабиевна</Преподаватель>
76     </Пара3-2>
77     <Пара4-1 type="dict">
78         <Время type="str">13:30-15:00</Время>
79         <Аудитория type="str">305</Аудитория>
80         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
81         <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>
82         <Преподаватель type="str">Яркеев Александр Сергеевич</Преподаватель>
83     </Пара4-1>
84     <Пара4-2 type="dict">
85         <Время type="str">13:30-15:00</Время>
86         <Аудитория type="str">303/2 (УСЛ)</Аудитория>
87         <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>
88         <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>
89         <Преподаватель type="str">Ткешелашвили Нино Мирабиевна</Преподаватель>
90     </Пара4-2>
91 </ЧетнаяНеделя>
92 </Понедельник>
93 </all>

```

Необязательное задание 1:

Для перевода использовал готовую библиотеку json2xml. Полученные результаты с помощью написанного парсера и готовой библиотеки полностью совпадают.

Необязательное задание 2:

Время десятикратного выполнения написанного парсера: 0.005594730377197266

Время десятикратного выполнения парсера готовой библиотеки: 0.0756826400756836

Написанный мною парсер переводит только необходимые для данного файла синтаксис, в отличии от парсера готовой библиотеки, который может перевести все, что есть в языке разметки. Поэтому мой парсер оказался быстрее.

Вывод: во время выполнения лабораторной работы изучил некоторые языки разметки, такие как XML, JSON, YAML, PROTOBUF; овладел навыками обработки данных на языке Python 3.x.