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. a) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.

b) Сравнить полученные результаты и объяснить их.

1. a) Используя свою программу и найденные готовые библиотеки, сравнить десятикратное время выполнения парсинга + конвертации в цикле.

b) Проанализировать полученные результаты и объяснить их.

**Листинг программы .py:**

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

**Содержание файла в исходном формате (JSON):**

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92 | {  "Понедельник" : {  "НечетнаяНеделя" : {  "Пара1-1" : {  "Время" : "08:20-09:50",  "Аудитория" : "303/2 (усл)",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Сорокин Роман Борисович"  },  "Пара1-2" : {  "Время" : "08:20-09:50",  "Аудитория" : "305",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Горбунов Михаил Витальевич"  },  "Пара2-1" : {  "Время" : "10:00-11:30",  "Аудитория" : "303/2 (усл)",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Сорокин Роман Борисович"  },  "Пара2-2" : {  "Время" : "10:00-11:30",  "Аудитория" : "305",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Горбунов Михаил Витальевич"  },  "Пара3" : {  "Время" : "11:40-13:10",  "Аудитория" : "369А",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ДИСКРЕТНАЯ МАТЕМАТИКА(ПРАК)",  "Преподаватель" : "Поляков Владимир Иванович"  },  "Пара4" : {  "Время" : "13:30-15:00",  "Аудитория" : "369А",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ДИСКРЕТНАЯ МАТЕМАТИКА(ПРАК)",  "Преподаватель" : "Поляков Владимир Иванович"  }  },  "ЧетнаяНеделя" : {  "Пара1" : {  "Время" : "08:20-09:50",  "Аудитория" : "303/2 (усл)",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ИНФОРМАТИКА(ЛАБ)",  "Преподаватель" : "Балакшин Павел Валерьевич"  },  "Пара2" : {  "Время" : "10:00-11:30",  "Аудитория" : "303/2 (усл)",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ИНФОРМАТИКА(ЛАБ)",  "Преподаватель" : "Балакшин Павел Валерьевич"  },  "Пара3-1" : {  "Время" : "11:40-13:10",  "Аудитория" : "305",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Яркеев Александр Сергеевич"  },  "Пара3-2" : {  "Время" : "11:40-13:10",  "Аудитория" : "303/2 (УСЛ)",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Ткешелашвили Нино Мирабиевна"  },  "Пара4-1" : {  "Время" : "13:30-15:00",  "Аудитория" : "305",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Яркеев Александр Сергеевич"  },  "Пара4-2" : {  "Время" : "13:30-15:00",  "Аудитория" : "303/2 (УСЛ)",  "Корпус" : "Кронверкский пр., д.49, лит.А",  "Предмет" : "ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ",  "Преподаватель" : "Ткешелашвили Нино Мирабиевна"  }  }  }  } |

**Содержание файла в результирующем формате (XML):**

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93 | <?xml version="1.0" ?>  <all>  <Понедельник type="dict">  <НечетнаяНеделя type="dict">  <Пара1-1 type="dict">  <Время type="str">08:20-09:50</Время>  <Аудитория type="str">303/2 (усл)</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Сорокин Роман Борисович</Преподаватель>  </Пара1-1>  <Пара1-2 type="dict">  <Время type="str">08:20-09:50</Время>  <Аудитория type="str">305</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Горбунов Михаил Витальевич</Преподаватель>  </Пара1-2>  <Пара2-1 type="dict">  <Время type="str">10:00-11:30</Время>  <Аудитория type="str">303/2 (усл)</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Сорокин Роман Борисович</Преподаватель>  </Пара2-1>  <Пара2-2 type="dict">  <Время type="str">10:00-11:30</Время>  <Аудитория type="str">305</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ПРОГРАММИРОВАНИЕ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Горбунов Михаил Витальевич</Преподаватель>  </Пара2-2>  <Пара3 type="dict">  <Время type="str">11:40-13:10</Время>  <Аудитория type="str">369А</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ДИСКРЕТНАЯ МАТЕМАТИКА(ПРАК)</Предмет>  <Преподаватель type="str">Поляков Владимир Иванович</Преподаватель>  </Пара3>  <Пара4 type="dict">  <Время type="str">13:30-15:00</Время>  <Аудитория type="str">369А</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ДИСКРЕТНАЯ МАТЕМАТИКА(ПРАК)</Предмет>  <Преподаватель type="str">Поляков Владимир Иванович</Преподаватель>  </Пара4>  </НечетнаяНеделя>  <ЧетнаяНеделя type="dict">  <Пара1 type="dict">  <Время type="str">08:20-09:50</Время>  <Аудитория type="str">303/2 (усл)</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ИНФОРМАТИКА(ЛАБ)</Предмет>  <Преподаватель type="str">Балакшин Павел Валерьевич</Преподаватель>  </Пара1>  <Пара2 type="dict">  <Время type="str">10:00-11:30</Время>  <Аудитория type="str">303/2 (усл)</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ИНФОРМАТИКА(ЛАБ)</Предмет>  <Преподаватель type="str">Балакшин Павел Валерьевич</Преподаватель>  </Пара2>  <Пара3-1 type="dict">  <Время type="str">11:40-13:10</Время>  <Аудитория type="str">305</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Яркеев Александр Сергеевич</Преподаватель>  </Пара3-1>  <Пара3-2 type="dict">  <Время type="str">11:40-13:10</Время>  <Аудитория type="str">303/2 (УСЛ)</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Ткешелашвили Нино Мирабиевна</Преподаватель>  </Пара3-2>  <Пара4-1 type="dict">  <Время type="str">13:30-15:00</Время>  <Аудитория type="str">305</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Яркеев Александр Сергеевич</Преподаватель>  </Пара4-1>  <Пара4-2 type="dict">  <Время type="str">13:30-15:00</Время>  <Аудитория type="str">303/2 (УСЛ)</Аудитория>  <Корпус type="str">Кронверкский пр., д.49, лит.А</Корпус>  <Предмет type="str">ОСНОВЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ(ЛАБ): 1/2 ГРУППЫ</Предмет>  <Преподаватель type="str">Ткешелашвили Нино Мирабиевна</Преподаватель>  </Пара4-2>  </ЧетнаяНеделя>  </Понедельник>  </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.