

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«Национальный исследовательский университет ИТМО»

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И
КОМПЬЮТЕРНОЙ ТЕХНИКИ

ЛАБОРАТОРНАЯ РАБОТА №4

по дисциплине

«ИНФОРМАТИКА»

Исследование протоколов,
форматов обмена информацией и языков разметки документов

Вариант 19

Выполнил:

Студент группы РЗ131

Новиков Даниил Дмитриевич

Проверил:

Авксентьева Елена Юрьевна, доцент

Содержание

Задание	3
Основные этапы вычисления	5
Заключение	9
Список литературы	10

Задание

1. Определить номер варианта как остаток деления на 36 последних двух цифр своего идентификационного номера в ISU. В случае, если в данный день недели нет занятий, то увеличить номер варианта на восемь.
2. Изучить форму Бэкуса-Наура.
3. Изучить основные принципы организации формальных грамматик.
4. Изучить особенности языков разметки/форматов JSON, YAML, XML.
5. Понять устройство страницы с расписанием на примере расписания лектора:
https://itmo.ru/ru/schedule/3/125598/raspisanie_zanyatiy.htm
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 45 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 путём простой замены метасимволов исходного формата на метасимволы результирующего формата. Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов.
8. **Дополнительное задание №1** (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
 - а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
 - б) Переписать исходный код, применив найденные библиотеки Регулярные выражения также нельзя использовать.
 - в) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.
9. **Дополнительное задание №2** (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
 - а) Переписать исходный код, добавив в него использование регулярных выражений.
 - б) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.
10. **Дополнительное задание №3** (позволяет набрать +25 процентов от максимального числа баллов БаРС за данную лабораторную).
 - а) Переписать исходный код таким образом, чтобы для решения задачи использовались формальные грамматики. То есть ваш код должен уметь осуществлять парсинг и конвертацию любых данных, представленных в исходном формате, в данные, представленные в результирующем формате: как с готовыми библиотеками и дополнительного задания №1.
 - б) Проверку осуществить как минимум для расписания с двумя учебными днями по два занятия в каждом.
 - в) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

11. **Дополнительное задание №4** (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).
- а) Используя свою исходную программу из обязательного задания и программы из дополнительных заданий, сравнить стократное время выполнения парсинга + конвертации в цикле.
 - б) Проанализировать полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.
12. **Дополнительное задание №5** (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).
- а) 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.
 - б) Проанализировать полученные результаты, объяснить особенности использования формата. Объяснение должно быть отражено в отчёте.

Основные этапы вычисления

Обязательное задание:

Исходный файл xml:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/timetable.xml>

Исходный код:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/task0.py>

Дополнительное задание 1:

Исходный код:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/task1.py>

Дополнительное задание 2:

Исходный код:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/task2.py>

Дополнительное задание 3:

Исходный код:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/task3.py>

Дополнительное задание 4:

Исходный код:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/task4.py>

Дополнительное задание 5:

Исходный код:

<https://github.com/Buratishkin/ITMO/blob/main/inf/lab4/code/task5.py>

Заключение

В ходе проделанной лабораторной работы, я познакомился с разными форматами файлов. Узнал о существовании формальных грамматик и БНФ. Написал парсеры, которые конвертируют json-файл в формат xml и tsv. Получил удовольствие от 5 часового кодинга рекурсивного парсера на формальных грамматиках.

Список литературы

1. [Электронный ресурс]: Википедия. Свободная энциклопедия. – Режим доступа: <https://ru.wikipedia.org/wiki/JSON>
2. Грошев А.С. Г89 Информатика: Учебник для вузов / А.С. Грошев. – Архангельск, Арханг. гос. техн. ун-т, 2010. - 470с. - Режим доступа: <http://arm.sies.uz/wp-content/uploads/2020/11/16-%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA-%D0%B0-2010-%D0%B4%D0%B0%D1%80%D1%81%D0%BB%D0%B8%D0%BA-%D0%93%D1%80%D0%BE%D1%88%D0%B5%D0%B2-%D0%90.%D0%A1.pdf>