«Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Направление подготовки: 09.03.01 - Информатика и вычислительная техника, Компьютерные системы и технологии

Дисциплина «Информатика»

Отчет по информатике

Лабораторная работа № 4

Вариант 18

Выполнил: Хабнер Г. Е. студент группы P3131

Проверил: Белозубов Александр Владимирович

г. Санкт-Петербург, 2022г.

Оглавление

[Задание 3](#_Toc119434008)

[Основные этапы вычисление 5](#_Toc119434009)

[Заключение 12](#_Toc119434010)

[Список литературы 13](#_Toc119434011)

# Задание

1. Определить номер варианта как остаток деления на 36 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.

4. Понять устройство страницы с расписанием для своей группы: <http://itmo.ru/ru/schedule/0/P3110/schedule.htm>

5. 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.

6. Обязательное задание (позволяет набрать до 65 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.

7. 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). a) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов. b) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать. c) Сравнить полученные результаты и объяснить их сходство/различие.

9. Дополнительное задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную). a) Переписать исходный код, добавив в него использование регулярных выражений. b) Сравнить полученные результаты и объяснить их сходство/различие.

10.Дополнительное задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную). a) Используя свою исходную программу из обязательного задания, программу из дополнительного задания №1 и программу из дополнительного задания №2, сравнить стократное время выполнения парсинга + конвертации в цикле. b) Проанализировать полученные результаты и объяснить их сходство/различие.

11.Дополнительное задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную). c) 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. d) Проанализировать полученные результаты, объяснить особенности использования формата.

12.Проверить, что все пункты задания выполнены и выполнены верно.

# Основные этапы вычисление

Задание 1



Задание 2

Сравнение решения с использованием библиотеки json2xml и нативного решения.

Нативное решение работает более быстро. В библиотеке в отличие от моего нативного решения учтены все нюансы перевода данных из json в xml и также обладает обширных функционалом для работы с данными.



Задание 3

Сравнение решения с использованием регулярных выражений и нативного решения.

Решение с использованием регулярных выражений выполняет конвертация намного медленнеe, из-за того, что использование регулярных выражений имеет смысл при работе с большими данными, но при этом, они упрощают поиск составления сложных выражений для поиска.



Задание 4

Сравнение скорости работы всех решений

* Время выполнения нативного решения: 0.01900076866149 сек.
* Время выполнения кода с использованием библиотеки json2xml: 0.574515104293 сек.
* Время выполнения кода с использованием регулярных выражений: 0.0949974060058 сек.

Нативное решение выигрывает по скорости, ведь в отличие от библиотеки, которая помимо конвертации данных предоставляет обширный интерфейс для работы с файлом, оно выполняет определенную задачу – конвертацию файла. Использование регулярных выражений выполняет поиск не особо эффективно так как они предусматривают работу с большими данными, а в нашем случае происходит лишь поиск тэгов и значений заключенных в них.



Задание 5

Конвертация файла в TOML

Для выполнения последнего задания я выбрал TOML. Он спроектирован для обеспечения человекочитаемости и однозначного преобразования в ассоциативный массив. TOML используется в ряде программных продуктов, поддерживается большим количеством языков программирования.



# Заключение

В ходе работы я познакомился с различными языками разметки и их особенностями. Узнал способы парсинга таких файлов и их конвертации из одного формата в другой.

# Список литературы

1. https://ru.wikipedia.org/wiki/Форма\_Бэкуса\_—\_Наура [В Интернете] // Форма Бэкуса — Наура. - 15 сентябрь 2021 г..
2. **В. П. Балакшин** https://web.telegram.org/d0248d44-3bcd-4c7e-8440-f7f1c823f26f [В Интернете] // "Презентация по информатике". - 2022 г..