НИУ ИТМО

Факультет программной инженерии и компьютерной техники

Отчет по Лабораторной работе №4    
по курсу “Информатика”

"Исследование протоколов, форматов обмена информацией и языков разметки документов"    
Вариант №22

Выполнила:

Студентка группы P3110

Бегинина Анастасия Алексеевна   
Преподаватели:

Балакшин Павел Валерьевич  
Рудникова Тамара Владимировна

Санкт-Петербург

2021

Оглавление

[Задания 3](#_Toc87391258)

[1. Обязательное задание 3](#_Toc87391259)

[2. Доп. задание No1 (+10 баллов) 3](#_Toc87391260)

[3. Доп. задание No2 (+10 баллов) 3](#_Toc87391261)

[4. Доп. задание No3 (+10 баллов) 3](#_Toc87391262)

[5. Доп. задание No4 (+5 баллов) 3](#_Toc87391263)

[Основные этапы работы 4](#_Toc87391264)

[Вывод 5](#_Toc87391265)

[Список литературы 6](#_Toc87391266)

[Приложения 7](#_Toc87391267)

# Задания

## 1. Обязательное задание

Написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.

## 2. Доп. задание No1 (+10 баллов)

1. Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
2. Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.
3. Сравнить полученные результаты и объяснить их сходство/различие.

## 3. Доп. задание No2 (+10 баллов)

a)  Переписать исходный код, добавив в него использование регулярных выражений.

b)  Сравнить полученные результаты и объяснить их сходство/различие.

## 4. Доп. задание No3 (+10 баллов)

a)  Используя свою исходную программу из обязательного задания, программу из дополнительного задания No1 и программу из дополнительного задания No2, сравнить десятикратное время выполнения парсинга + конвертации в цикле.

b)  Проанализировать полученные результаты и объяснить их сходство/различие.

## 5. Доп. задание No4 (+5 баллов)

a)  Переписать исходную, чтобы она осуществляла парсинг и конвертацию исходного файла в любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.

b)  Проанализировать полученные результаты, объяснить осоебнности использованного формата.

# Основные этапы работы

1. Определение варианта

Мой вариант 335038 % 36 = 22. Перевести из YAML в JSON файл с расписанием в четверг.

1. Изначально составленный файл YAML

(см. Приложение 1)

1. Весь исходный код и json-файлы см. <https://github.com/nastyabeggin/itmo_computer_science>
2. Различие json-файлов из доп. задания 1 и из основного задания:

Файлы отличаются только порядком характеристик, так как библиотека автоматически сортирует по алфавиту. (см. Приложения 2, 3)

1. Различие json-файлов из доп. задания 2 и из основного задания:  
   Файлы абсолютно идентичны.
2. Сравнение времени  
   Сравним время выполнения программ с помощью метода motonic\_ns() из библиотеки time.

Результат:

0.00332875 seconds basic

0.03517666 seconds using libs

0.04349375 seconds with regex

Базовое решение работает быстрее всего, с библиотеками дольше, так как нужно дополнительное время на их загрузку. Дольше всего работает решение с использованием регулярных выражений, так как методы из библиотеки re вызываются больше всего раз

1. CSV-формат  
   Особенности данного формата: предназначен для представления табличных данных. В качестве разделителя по умолчанию используются запятые. (см. Приложение 4)

# Вывод

В ходе выполнения лабораторной работы я научилась работать с JSON, YAML, CSV файлами. Укрепила навык использования регулярных выражений и сторонних библиотек.

# Список литературы

1. Регулярные выражения в Python от простого к сложному. Подробности, примеры, картинки, упражнения // habr.com: блог об IT. 2007 URL: https://habr.com/ru/post/349860/ (дата обращения: 9.11.2021).
2. YAML vs JSON. // habr.com: блог об IT. 2007 URL: https://habr.com/ru/company/rambler\_and\_co/blog/525498/ (дата обращения: 9.11.2021).

# Приложения

Приложение 1.

Text

Description automatically generated

Приложение 2.

Text

Description automatically generated

Приложение 3.

Text

Description automatically generated

Приложение 4.

