Санкт-Петербургский Национальный Исследовательский Университет ИТМО

Факультет программной инженерии и компьютерной техники

Лабораторная работа №4

“Исследование протоколов,

форматов обмена информацией и языков разметки документов”

По дисциплине

“Информатика”

Вариант 14

Выполнил:

Студент группы P3117

Пономарёв М. И.

Преподаватель:

Машина Е. А.



Оглавление

[Задание 3](#_Toc119018960)

[Основные этапы вычисления 4](#_Toc119018961)

[getHTMLwithBs4.py 4](#_Toc119018962)

[HtmlToJson.py 4](#_Toc119018963)

[JsonToYaml.py 5](#_Toc119018964)

[JsonToYamlLibrary.py 6](#_Toc119018965)

[JsonToYamlRegex.py 6](#_Toc119018966)

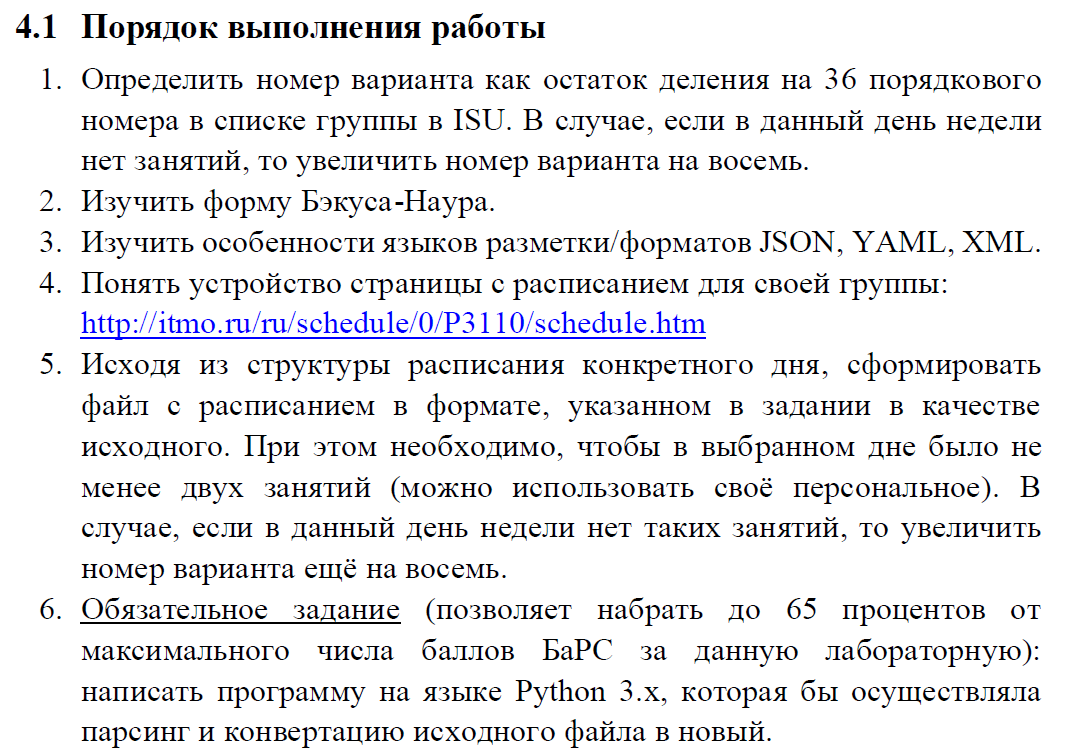
[main.py 7](#_Toc119018967)

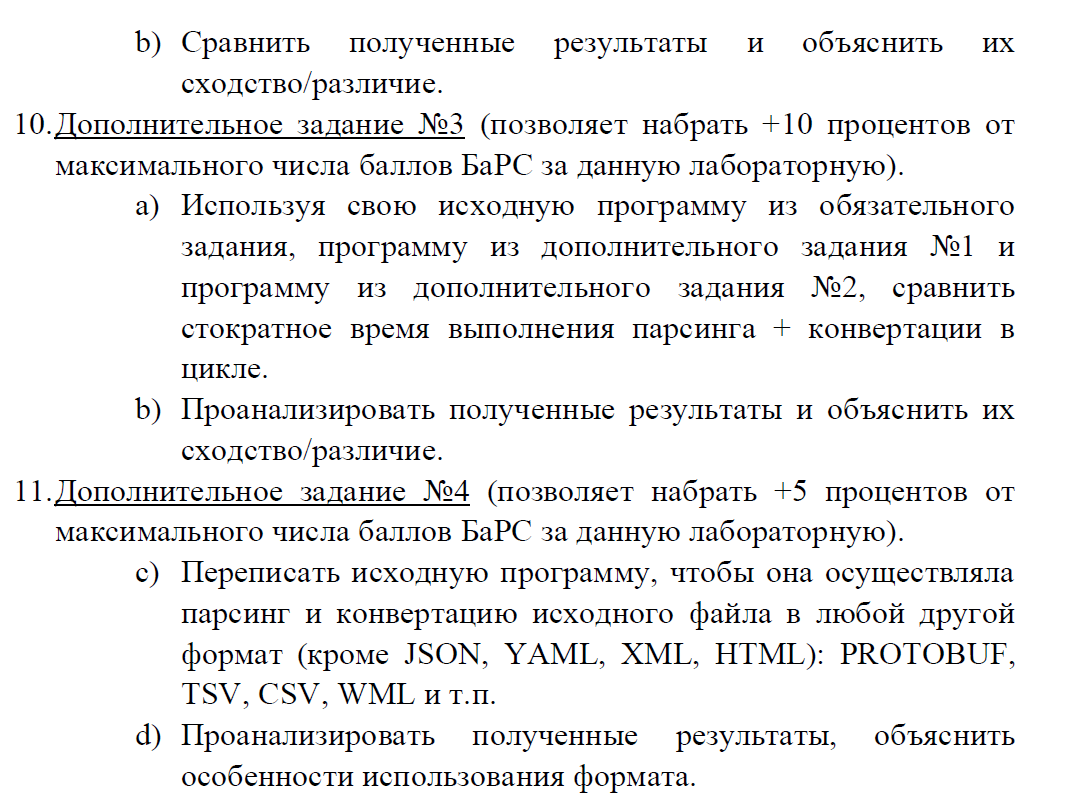
[compare.py 7](#_Toc119018968)

[Вывод 8](#_Toc119018969)

Задание

Изображение выглядит как текст

Автоматически созданное описание



Основные этапы вычисления

## getHTMLwithBs4.py



## HtmlToJson.py



## JsonToYaml.py



## JsonToYamlLibrary.py



## JsonToYamlRegex.py



## main.py



## compare.py



Результат выполнения программы compare.py:

|  |
| --- |
| Парсинг через библиотеку: 4.575491189956665 seconds  Парсинг написанной функцией: 0.23781204223632812 seconds  Разница составляет ~1923% или 4.337679147720337 секунд |

Результаты работы main.py содержаться в директории results/

Вывод

Во время выполнения работы я использовал библиотеки json, requests, bs4, yaml для перевода файлов между заданными форматами. Наглядно заметил читаемость и удобство формата yaml в сравнении с json.