

Университет ИТМО

Факультет программной инженерии и компьютерной техники
Образовательная программа системное и прикладное
программное обеспечение

Лабораторная работа №4
По дисциплине "Информатика"
Вариант 26

Выполнил студент группы Р3109
Евграфов Артём Андреевич
Проверил:
Рыбаков Степан Дмитриевич

Санкт-Петербург 2024

Содержание

1. Задание	2
2. Исходные файлы	2
3. Обязательное задание	2
4. Дополнительное задание №1	3
5. Дополнительное задание №2	3
6. Дополнительное задание №4	3
7. Дополнительное задание №5	3
8. Вывод	3
9. Список литературы	4

1. Задание

Обязательное задание (позволяет набрать до 45 процентов от максимального числа баллов БаРС за данную лабораторную): написать программу на языке Python 3.x или любом другом, 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.

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. Объяснение должно быть отражено в отчёте.

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. Объяснение должно быть отражено в отчёте.

Дополнительное задание №3 (позволяет набрать +25 процентов от максимального числа баллов БаРС за данную лабораторную).

- a) 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 из дополнительного задания №1.
- b) Проверку осуществить как минимум для расписания с двумя учебными днями по два занятия в каждом.
- c) Сравнить полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

Дополнительное задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).

- a) Используя свою исходную программу из обязательного задания и программы из дополнительных заданий, сравнить стократное время выполнения парсинга + конвертации в цикле.
- b) Проанализировать полученные результаты и объяснить их сходство/различие. Объяснение должно быть отражено в отчёте.

Дополнительное задание №5 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).

- a) 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.
- b) Проанализировать полученные результаты, объяснить особенности использования формата. Объяснение должно быть отражено в отчёте.

2. Исходные файлы

[JSON файл](#)

[YAML файл](#)

[TOML файл](#)

3. Обязательное задание

Решение задания по [ссылке](#)

JSON (JavaScript Object Notation) — это текстовый формат, используемый для представления структурированных данных. JSON легко читается человеком и одновременно легко обрабатывается компьютерами.

YAML (YAML Ain't Markup Language) — это текстовый формат, предназначенный для сериализации

(процесс преобразования данных или объектов в формат, который можно легко сохранить или передать, а затем восстановить в исходном виде) данных. YAML разработан для удобства человека, чтобы писать и читать конфигурационные файлы и обмениваться данными между системами.

4. Дополнительное задание №1

Решение задания по [ссылке](#)

Парсинг выполнен с помощью библиотек json и yaml. Различий между результатом работы программы предыдущего задания и текущего нет, поскольку оба парсера для данного JSON файла полностью эквивалентны.

5. Дополнительное задание №2

Решение задания по [ссылке](#)

Различий между результатом работы программы предыдущего задания и текущего нет, поскольку множество строк подходящих под условия (if-ы) из обязательного задания совпадает с множеством строк, подходящих под регулярные выражения из данного задания.

6. Дополнительное задание №4

Решение задания по [ссылке](#)

```
C:\Users\minec\AppData\Local\Programs\Python\Python313\python.exe C:/Users/minec/PycharmProjects/minec/task4.py
0.07585050002671778 seconds - required task
0.0824700000230223 seconds - 1 task
0.08233889995608479 seconds - 2 task
```

Рис. 1: Результат работы программы

Дольше всего работает программа с готовыми библиотеками, так как она реализует полноценный парсинг с полной обработкой входных и выходных файлов. Код с регулярными выражениями работает немного дольше основного кода, поскольку сами регулярные выражения довольно медленные, однако разница незначительна из-за малых размеров читаемого файла.

7. Дополнительное задание №5

Парсер из JSON в TOML доступен по [ссылке](#)

TOML (Tom's Obvious, Minimal Language) — это язык разметки, который используется для настройки и конфигурации программ. Его основное назначение — предоставлять простой, человекочитаемый формат для структурированных данных. TOML особенно популярен в средах, где требуется легко настраивать параметры приложений или передавать конфигурации между системами.

Конфигурационные файлы — это текстовые файлы, которые используются для настройки программного обеспечения или операционных систем. Они позволяют разработчикам и пользователям задавать параметры работы приложения, не изменяя его исходный код.

8. Вывод

Во время выполнения лабораторной работы я узнал о языках разметки JSON, TOML и YAML, научился с ними работать и переводить один тип данных в другой с помощью встроенных средств языка Python. Научился работать с библиотекой timeit для измерения времени и некоторыми библиотеками для анализа этих форматов и автоматического парсинга между ними.

9. Список литературы

Балакшин П.В., Соснин В.В., Калинин И.В., Малышева Т.А., Раков С.В., Рущенко Н.Г., Дергачев А.М. Информатика: лабораторные работы и тесты: Учебно-методическое пособие / Рецензент: Поляков В.И. - Санкт-Петербург: Университет ИТМО, 2019. - 56 с. - экз. - Режим доступа: https://books.ifmo.ru/catalog/2019/catalog_2024.htm

Лямин А.В., Череповская Е.Н. Объектно-ориентированное программирование. Компьютерный практикум. – СПб: Университет ИТМО, 2017. – 143 с. – Режим доступа: <https://books.ifmo.ru/file/pdf/2256.pdf>