

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский университет ИТМО»
(Университет ИТМО)

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И
КОМПЬЮТЕРНОЙ ТЕХНИКИ

ЛАБОРАТОРНАЯ РАБОТА №2

по дисциплине
«ИНФОРМАТИКА»

по теме:

ИССЛЕДОВАНИЕ ПРОТОКОЛОВ, ФОРМАТОВ ОБМЕНА
ИНФОРМАЦИЕЙ И ЯЗЫКОВ РАЗМЕТКИ ДОКУМЕНТОВ

Вариант №13

Выполнил:

Студент группы Р3117

Павленко И. Д.

Проверил:

Марухленко Д. С.

Санкт-Петербург 2024

СОДЕРЖАНИЕ

1	ЗАДАНИЕ	3
2	ОСНОВНЫЕ ЭТАПЫ РЕШЕНИЯ	5
2.1	Обязательное задание	5
2.2	Дополнительное задание №1	5
2.3	Дополнительное задание №2	5
2.4	Дополнительное задание №3	5
2.5	Дополнительное задание №4	5
2.6	Дополнительное задание №5	6
3	ЗАКЛЮЧЕНИЕ	7
	СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	8

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). а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов. б) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать. с) 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). 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.

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 из дополнительного задания №1. б) Проверку осуществить как минимум для расписания с двумя учебными днями по два занятия в каждом. с) 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). а) Используя свою исходную программу из обязательного задания и программы из дополнитель-

ных заданий, сравнить стократное время выполнения парсинга + конвертации в цикле. б) 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). 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.

2 ОСНОВНЫЕ ЭТАПЫ РЕШЕНИЯ

2.1 Обязательное задание

Код программы находится на [Github](#).

2.2 Дополнительное задание №1

Код программы находится на [Github](#).

Различий в выводе программ нет. Очень хорошо видно уменьшение кода, необходимого для конвертации файла.

2.3 Дополнительное задание №2

Код программы находится на [Github](#)

Различий в выводе программ нет. Регулярные выражения используются для "отлавливания" тегов и значений этих тегов. Логика конвертации не изменена.

2.4 Дополнительное задание №3

Код программы находится на [Github](#).

Различий в выводе программ нет. Код обязательного задания изначально писался с использованием формальных грамматик.

2.5 Дополнительное задание №4

Код программы находится на [Github](#)

Обязательное задание и №3: 0.022017240524291992

Дополнительное задание №1: 0.03467750549316406

Дополнительное задание №2: 0.027050495147705078

Программа первого доп. задания выполняется дольше из-за того, что программа обязательного задания выполняется "in-place".

2.6 Дополнительное задание №5

Код программы находится на [Github](#).

CSV - простой текстовый формат для хранения табличных данных, где значения разделяются запятыми (или другими символами-разделителями). Формат CSV удобен для работы с большими объемами данных и легко читается как человеком, так и машиной.

3 ЗАКЛЮЧЕНИЕ

В ходе выполнения работы я узнал о языках разметки XML и JSON. Научился с ними работать, а также конвертировать XML в JSON.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *А.С. Грошев.* Информатика: Учебник для вузов. — Архангельский Государственный технический Университет, 2016. — С. 470.
2. *А.М. Балакишин П.В. Соснин В.В. Калинин И.В. Малышева Т.А. Раков С.В. Рущенко Н.Г. Дергачев.* Информатика: лабораторные работы и тесты: учебно-методическое пособие. — Университет ИТМО, 2019. — С. 56.