Федеральное государственное автономное образовательное учреждение высшего образования **«Национальный исследовательский университет ИТМО»**

Факультет Программной Инженерии и Компьютерной Техники

**Лабораторная работа по информатике №4**

вариант: 11

Преподаватель: Малышева Татьяна Алексеевна

Выполнил: Дагаев Даниил Сергеевич

Группа: Р3108

Санкт-Петербург, 2021г

Оглавление

[Задание 2](#_Toc89125223)

[Ход работы: 2](#_Toc89125224)

[Обязательное задание 3](#_Toc89125225)

[Дополнительное задание №1 4](#_Toc89125226)

[Вывод 5](#_Toc89125227)

[Список литературы. 5](#_Toc89125228)

# Задание



* 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.
* Дополнительное задание задание №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
  + Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
  + Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.
  + Сравнить полученные результаты и объяснить их сходство/различие.
* Дополнительное задание задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
  + Переписать исходный код, добавив в него использование регулярных выражений.
  + Сравнить полученные результаты и объяснить их сходство/различие.
* Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
  + Используя свою исходную программу из обязательного задания, программу из дополнительного задания №1 и программу из дополнительного задания №2, сравнить десятикратное время выполнения парсинга + конвертации в цикле.
  + Проанализировать полученные результаты и объяснить их сходство/различие.
* Дополнительное задание задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную.
  + 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.
  + Проанализировать полученные результаты, объяснить особенности использованного формата.

# Ход работы:

339923 % 36 = 11



# Обязательное задание

import timeit

def Parse():

r = open("lab3.xml", 'r', encoding="utf-8")

w = open("YAML.yaml", 'w', encoding="utf-8")

listXml = r.readlines()

countTab = 0;

for text in listXml:

ctext = text

while len(ctext) != 0:

if ctext[0] == ' ':

ctext = ctext[1:]

elif ctext[0] == '\n' or ctext[0] == '\t':

ctext = ctext[1:]

elif "</" in ctext[:2]:

countTab -= 1

l = ctext.find(">")

ctext = ctext[l+1:]

elif "<" in ctext[0] and ctext.count('>') == 2:

f = ctext.find("<")

l = ctext.find(">")

w.write( countTab \* ' ' + ctext[f+1:l] + ":")

ctext = ctext[l+1:]

countTab += 1

elif "<" in ctext[0]:

f = ctext.find("<")

l = ctext.find(">")

w.write( countTab \* ' ' + ctext[f+1:l] + ":" + '\n')

ctext = ctext[l+1:]

countTab += 1

else:

if "<" in ctext:

f = ctext.find("<")

if ":" in ctext[:f] or ctext[:f].isdigit():

w.write( ' ' + "'" + ctext[:f] + "'" + '\n')

ctext = ctext[f:]

else:

w.write(' ' + ctext[:f] + '\n')

ctext = ctext[f:]

else:

if ":" in ctext or ctext.isdigit():

w.write(' ' + "'" + ctext + "'")

ctext = ctext[len(ctext):]

else:

w.write(' ' + ctext)

ctext = ctext[len(ctext):]

Parse()

print('Я надеюсь это работает')

print(timeit.timeit(Parse,number=1))

**Результат:**

0.0034380999999999995

Файл YAML.yaml:

timetable:

day: Ср

lesson1:

time: '11:40-11:30'

room: 369A

subject: Дискретная Математика(практика)

teacher: Поляков Владимир Иванович

location: Кронверкскийпр., д.49, лит.А

parity: false

format: Очно-дистанционный

weeks: 3, 5, 7, 9, 11, 13, 15, 17

lesson2:

time: '11:40-11:30'

room: 369A

subject: Дискретная Математика(практика)

teacher: Поляков Владимир Иванович

location: Кронверкскийпр., д.49, лит.А

parity: false

format: Очно-дистанционный

weeks: 3, 5, 7, 9, 11, 13, 15, 17

# Дополнительное задание №1

import xmlplain

import timeit

def Parse():

with open("lab3.xml", "r", encoding="utf8") as read:

root = xmlplain.xml\_to\_obj(read, strip\_space=True, fold\_dict=True)

with open("YAML2.yaml", "w", encoding="utf8") as out:

xmlplain.obj\_to\_yaml(root, out)

print(timeit.timeit(Parse,number=1))

**Результат:**

0.0034380999999999995

Файл YAML2.yaml:

timetable:

day: Ср

lesson1:

time: 11:40-11:30

room: 369A

subject: Дискретная Математика(практика)

teacher: Поляков Владимир Иванович

location: Кронверкскийпр., д.49, лит.А

parity: 'false'

format: Очно-дистанционный

weeks: 3, 5, 7, 9, 11, 13, 15, 17

lesson2:

time: 11:40-11:30

room: 369A

subject: Дискретная Математика(практика)

teacher: Поляков Владимир Иванович

location: Кронверкскийпр., д.49, лит.А

parity: 'false'

format: Очно-дистанционный

weeks: 3, 5, 7, 9, 11, 13, 15, 17

# Вывод

Во время выполнения лабораторной работы я узнал о языках разметки XML и YAML, научился с ними работать и переводить один языках разметки в другой помощью средств языка Python. Научился работать с библиотекой timeit для измерения времени, и некоторыми библиотеками для автоматического парсинга.

# Список литературы.

1. 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/book/2248/informatika:_laboratornye_raboty_i_testy:_uchebno-metodicheskoe_posobie_/_recenzent:_polyakov_v.i..htm>

1. ГрошевА.С. Г89 Информатика: Учебник для вузов / А.С. Грошев. – Архангельск, Арханг. гос. техн. ун-т, 2010. -470с. -Режим доступа <https://narfu.ru/university/library/books/0690.pdf>