Министерство образования Республики Беларусь

Учреждение образования

“Брестский государственный технический университет”

Кафедра интеллектуально-информационных технологий

Лабораторная работа №3

По дисциплине «Языковые процессы интеллектуальных систем»

Выполнил:

студент 4 курса

группы ИИ-22

Копанчук Е. Р.

Проверил:

Монтик Н. С.

Брест-2024

**Цель работы:** изучить основных понятий теории регулярных грамматик, ознакомиться с назначением и принципами работы лексических анализаторов (сканеров), получить практические навыки построения сканера на примере заданного простейшего входного языка.

Написать программу для решения задачи согласно варианту. Данная лабораторная работа аналогична лабораторной работе №2, но при разработке должны использоваться регулярные выражения.

Вариант лабораторной работы соответствует варианту лабораторной работы №2.

**Задание**

9. Входной язык содержит последовательность описаний массивов в соответствии со спецификацией языка Паскаль, разделенных символами ; (точка с запятой). Считать, что массивы могут содержать только элементы скалярных типов integer, real, byte, word и char.

**Код программы:**

import re

def determine\_token\_type(token):

if re.match(r'^[0-9]+$', token):

return 'Integer'

elif re.match(r'^[0-9]\*\.[0-9]+$', token):

return 'Real'

elif re.match(r"^'[a-zA-Z]$|^\"[a-zA-Z]$", token):

return 'Char'

elif re.match(r'^[a-zA-Zа-яА-Я\_][a-zA-Zа-яА-Я0-9\_]\*$', token):

return 'Identifier'

elif token in ['+', '-', '\*', '/', '=', '==', '!', '&&', '||', '(', ')', '[', ']', '{', '}', ',', ';', '.', """'"""]:

return 'Operator'

elif token.startswith('//'):

return 'Comment'

else:

return 'Unknown'

def lexical\_analysis(text):

tokens = re.findall(r'\b\w+\b|[\'"\[\]\{\}\(\),;\.\+\-\\*/=\<\>\^]|==|!=|<=|>=|&&|\|\|', text)

token\_table = []

errors = []

for token in tokens:

token\_type = determine\_token\_type(token)

if token\_type == 'Unknown':

errors.append(f"Unknown token: '{token}'")

else:

token\_table.append((token, token\_type))

return token\_table, errors

file\_path = "D:/7 семестр/ЯПИС лабы/ЯПИС лаба №2/Массивы.txt"

with open(file\_path, 'r', encoding='utf-8') as file:

file\_data = file.read()

token\_table, errors = lexical\_analysis(file\_data)

print("Таблица лексем:")

for token, token\_type in token\_table:

print(f"Токен: '{token}', Тип: {token\_type}")

if errors:

print("\nОшибки:")

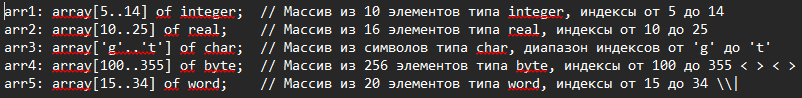
for error in errors:

print(error)

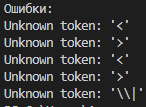
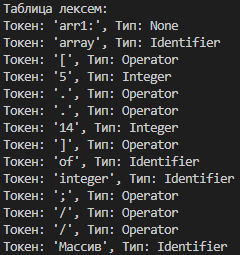
else:

print("\nОшибок не обнаружено.")

**Содержимое текстового файла:**



**Вывод программы:**



**Вывод:** изучил основные понятия теории регулярных грамматик, ознакомился с назначением и принципами работы лексических анализаторов (сканеров), получил практические навыки построения сканера на примере заданного простейшего входного языка.