Министерство образования Республики Беларусь Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»

Факультет компьютерных систем и сетей Кафедра информатики Дисциплина «Методы трансляции»

ОТЧЕТ

К лабораторной работе № 5 на тему «Интерпретация исходного кода»

Выполнил Е. А. Киселёва

Проверил Н. Ю. Гриценко

СОДЕРЖАНИЕ

1 Постановка задачи	. 3
2 Краткие теоретические сведения	
3 Результаты выполнения лабораторной работы	
Выводы	
Список использованных источников	. 8
Приложение А (обязательное) Листинг исходного кода	. 9

1 ПОСТАНОВКА ЗАДАЧИ

Целью выполнения данной лабораторной работы является на основе результатов анализа лабораторных работы 1-4 выполнить трансляцию программы с языка программирования С++ на язык программирования Python, после чего выполнить интерпретацию программы.

2 КРАТКИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

К этапам трансляции относятся следующие этапы:

- лексический анализ;
- синтаксический анализ;
- семантический анализ;
- оптимизация;
- генерация кода.

На этапе генерации компилятор создает код, который представляет собой набор инструкций, понятных для целевой аппаратной платформы, итоговый файл компилируется в исполняемый файл, который может быть запущен на целевой платформе без необходимости наличия кода.

Фаза эмуляции интерпретатора происходит во время выполнения программы. В отличие от компилятора, интерпретатор работает с кодом напрямую, без предварительной генерации машинного кода.

Лексический анализатор — первый этап трансляции. Лексический анализатор читает поток символов, составляющих исходную программу, и группирует эти символы в лексемы или значащие последовательности. Лексема — это элементарная единица, которая может являться ключевым словом, идентификатором, константным значением. Для каждой лексемы анализатор строит токен, который по сути является кортежем, содержащим имя и значение.[1]

Синтаксический анализатор выясняет, удовлетворяют ли предложения, из которых состоит исходная программа, правилам грамматики языка программирования. Синтаксический анализатор получает на вход результат лексического анализатора и разбирает его в соответствии с грамматикой. Результат синтаксического анализа обычно представляется в виде синтаксического дерева разбора.[2]

Семантический анализ обычно заключается в проверке правильности типа и вида всех идентификаторов и данных, используемых в программе.

Семантический анализатор использует синтаксическое дерево и информацию из таблицы символов для проверки исходной программы на семантическую согласованность с определением языка. Он также собирает информацию о типах и сохраняет ее в синтаксическом дереве или в таблице идентификаторов для последующего использования в процессе генерации промежуточного кода.

В данной лабораторной работе были использованы результаты анализа лексического, синтаксического и семантического анализаторов, после чего каждый узел дерева разбора был переведен с языка программирования С++ на язык программирования Руthon. После чего была выполнена интерпретация программ. Программами называются тестовые исходные коды, представленные в лабораторной работе 1.

З РЕЗУЛЬТАТЫ ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ

В ходе лабораторной работы был реализован транслятор программ с языка программирования С++ на язык программирования Python с последующей интерпретацией кода.

Листинг первой части тестового кода, представляющего собой быструю сортировку, представлен на рисунке 3.1.

```
include <iostream>
void quickSort(int *array, int first, int last)
   mid = array[(f + l) / 2];
       while (array[l] > mid) { l--; }
   if (first < l)</pre>
void printArray(int *arr, int size)
       cout << arr[i] << " ";
```

Рисунок 3.1 – Листинг первой части тестового кода

Листинг второй части тестового кода представлен на рисунке 3.2.

```
void printArray(int *arr, int size)

{
    for (int i = 0; i < size; i++)
    {
        cout << arr[i] << " ";
    }
    cout << endl;

arr[0] = 64;
    arr[1] = -322;
    arr[2] = 10;
    arr[3] = 22;
    arr[4] = -1;
    arr[6] = 4;
    arr[6] = 100;
    arr[7] = 100;
    arr[8] = 21;

    n = 9;

    cout << "ESKAA: \n";
    printArray(arr, n);

    quickSort(arr, 0, n - 1);
    cout << "\nSorted Array: \n";
    printArray(arr, n);

return 0;

arr 0;</pre>
```

Рисунок 3.2 – Листинг второй части тестового кода

Результат интерпретации исходного кода представлен на рисунке 3.3.

```
ESKAA:
64 -322 10 22 -1 4 100 100 21

Sorted Array:
-322 -1 4 10 21 22 64 100 100
```

Рисунок 3.3 – Результат интерпретации исходного кода

Таким образом в ходе лабораторной работы был реализован интерпретатор для программ на языке C++, который переводит их на язык программирования Python после чего проводит интерпретацию полученного при трансляции кода.

выводы

В ходе лабораторной работы был реализован был реализован интерпретатор для программ на языке C++, который переводит их на язык программирования Python после чего проводит интерпретацию полученного при трансляции кода.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- [1] Лексический анализатор [Электронный ресурс]. Режим доступа: https://csc.sibsutis.ru/sites/csc.sibsutis.ru/files/courses/trans/. Дата доступа: 27.02.2024.
- [2] Синтаксический анализатор [Электронный ресурс]. Режим доступа: https://csc.sibsutis.ru/sites/csc.sibsutis.ru/files/courses/trans/. Дата доступа: 27.02.2024.
- [3] Введение в C++ [Электронный ресурс]. Режим доступа: https://metanit.com/cpp/tutorial/2.5.php. Дата доступа: 28.02.2024.
- [4] Типы данных [Электронный ресурс]. Режим доступа: https://metanit.com/cpp/tutorial/2.3.php. Дата доступа: 28.02.2024.
- [5] Операторы в C++ [Электронный ресурс]. Режим доступа: https://learn.microsoft.com/ru-ru/cpp/c-operators. Дата доступа: 27.02.2024.
- [6] Функции С++ [Электронный ресурс]. Режим доступа: https://metanit.com/cpp/tutorial/3.1.php. Дата доступа: 27.02.2024.
- [7] Классы С++ [Электронный ресурс]. Режим доступа: https://ravesli.com/urok-113-klassy-obekty-i-metody-klassov/. Дата доступа: 27.02.2024.

приложение а

(обязательное)

Листинг исходного кода

Листинг 1 – Программный код parser.py

import types

```
from core.checks import *
from core.tree import *
def CreateFunctionObj(name, args, code):
       Creates a new function object from the given data.
    # Build function signature string from arg types
    args_str = ", ".join(args)
function_str = f"def {name}({args_str}):\n{code}"
   print(function str)
    # Compile function
    compiled func = compile(function str, "<string>", "exec")
    func code = next((c for c in compiled func.co consts if isinstance(c,
types.CodeType)), None)
    # Creation error handler
    if func code is None:
        raise ValueError("Unable to find function code object.")
    return types.FunctionType(func code, globals(), name)
class Translator:
       CPP to Python translator.
    def init (self, tree root, literal table, variable table):
            Initializes the translator object which can parse CPP AST and
translate it to Python.
        # Tree provided to translation
        self.Tree = tree root
        # Environment, constants and variables
        self.LiteralTable = literal table
        self.VariableTable = variable table
        self.Translate()
    def Translate(self):
           Core function of the translator: parses the function from AST and
translates them to python in order to use.
```

```
if self. Tree is None:
            return
        children = self.Tree.GetChildren()
        for child in children:
            if child.Type == SyntaxTreNodeTypes.FUNCTION DECLARATION:
                func nodes = child.GetChildren()
                # Get function name
                func name =
self.GetVariable(func nodes[1].GetLexeme().itemValue).itemName
                # Get function arguments
                func args = self.GetFunctionArguments(func nodes[2])
                # Get function body code
                body instructions nodes = func nodes[3].GetChildren()
                instructions = []
                for instruction node in body instructions nodes:
                    if not instruction node:
                    instruction = self.ParseInstruction(instruction node, 1)
                    if instruction is not None:
                        instructions.append(instruction)
                func_code = "\n".join(instructions)
                # Get function object
                function obj = CreateFunctionObj(func name, func args,
func code)
                # Assign function object to global scope
                globals()[func name] = function obj
        # Call main function
        globals()["main"]()
    def GetFunctionArguments(self, node):
            Parses function arguments node to get arguments in 'Python' form.
        arguments = []
        for arg in node.GetChildren():
            arg children = arg.GetChildren()
self.GetVariable(arg children[1].GetLexeme().itemValue)
            arg type = None
            if isinstance(variable.itemType, list):
               arg_type = "list"
            else:
                arg type = self.GetArgType(variable.itemType)
            if arg type is not None:
                arg name = variable.itemName
                arguments.append(f"{arg name}: {arg type}")
            else:
```

```
raise ValueError("Function argument type error")
        return arguments
    def GetArgType(self, arg type):
           Parses arguments CPP type to Python type.
        if arg type == Language.VariableTypes.INT:
            return "int"
        elif arg type == Language.VariableTypes.STRING:
            return "str"
        elif arg type == Language.VariableTypes.BOOL:
            return "bool"
        elif arg type == Language.VariableTypes.DOUBLE:
            return "float"
    def GetVariable(self, variable id):
           Gets variable item from the table.
        variable = None
            variable = [v for v in self.VariableTable if v.itemId ==
variable id][0]
        except:
            raise ValueError("Bad variable id")
        return variable
    def GetLiteral(self, literal id):
           Gets literal item from the table.
        literal = None
            literal = [v for v in self.LiteralTable.Literals if v.itemId ==
literal id][0]
        except:
            raise ValueError("Bad variable id")
        return literal
    def ParseInstruction(self, instruction node, level):
            Parses given instruction node.
        instruction = None
        lexeme = None
        cur level = '\t' * level
        if instruction node.Type == SyntaxTreNodeTypes.COMMON:
            lexeme = instruction node.GetLexeme()
        if instruction node.Type == SyntaxTreNodeTypes.DECLARATION:
```

```
instruction =
str(self.ParseVariableDeclarationStatement(instruction node, level))
        elif instruction node. Type == SyntaxTreNodeTypes. FUNCTION CALL:
            instruction =
str(self.ParseFunctionCallStatement(instruction node, level))
        elif lexeme and lexeme.itemValue == Language.Operators.EQUAL:
            instruction = str(self.ParseOperator(instruction node, level))
        elif lexeme and lexeme.itemValue == Language.KeyWords.IF:
            instruction = str(self.ParseIfStatement(instruction node, level))
        elif lexeme and lexeme.itemValue == Language.KeyWords.WHILE:
            instruction = str(self.ParseWhileStatement(instruction node,
level))
        elif lexeme and lexeme.itemValue == Language.KeyWords.FOR:
            instruction = str(self.ParseForStatement(instruction node,
level))
        elif lexeme and lexeme.itemValue == Language.KeyWords.DO:
            instruction = str(self.ParseDoWhileStatement(instruction node,
level))
        elif lexeme and lexeme.itemValue in [Language.Operators.INCREMENT,
Language.Operators.DECREMENT]:
            instruction =
str(self.ParseUnaryOperatorStatement(instruction node, level))
        elif lexeme and lexeme.itemValue == Language.KeyWords.CIN:
            instruction = str(self.ParseCinStatement(instruction node,
        elif lexeme and lexeme.itemValue == Language.KeyWords.COUT:
            instruction = str(self.ParseCoutStatement(instruction node,
level))
        elif lexeme and lexeme.itemValue in [Language.KeyWords.RETURN,
Language.KeyWords.EXIT]:
            instruction = str(self.ParseReturnExitStatement(instruction node,
level))
        elif lexeme and lexeme.itemValue == Language.KeyWords.BREAK:
            instruction = f"{cur level}break"
        elif lexeme and lexeme.itemValue == Language.KeyWords.CONTINUE:
            instruction = f"{cur level}continue"
        if instruction is not None:
            return instruction
    def ParseOperator(self, operator node, level):
            Parses the operator expression to Python.
        parts = []
        operator = None
        current level = level * '\t'
        if operator node.GetLexeme().itemType ==
Language.LexemeTypes.IDENTIFIER:
            return
str(self.GetVariable(operator node.GetLexeme().itemValue).itemName)
        elif operator node.GetLexeme().itemType ==
Language.LexemeTypes.INT NUM:
            return
str(self.GetLiteral(operator node.GetLexeme().itemValue).itemValue)
        try:
            operator =
inverted operators[operator node.GetLexeme().itemValue]
        except:
```

```
raise ValueError("Unknown operator")
        if operator == "-" and len(operator node.GetChildren()) == 1:
            node = operator node.GetChildren()[0]
            lexeme = operator node.GetChildren()[0].GetLexeme()
            if not node.GetChildren():
                if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                    return "-" +
str(self.GetVariable(lexeme.itemValue).itemName)
                elif lexeme.itemType in [Language.LexemeTypes.INT NUM,
Language.LexemeTypes.DOUBLE NUM]:
                    return "-" +
str(self.GetLiteral(lexeme.itemValue).itemValue)
            elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                var name = str(self.GetVariable(lexeme.itemValue).itemName)
                return "-" +
f"{var name}[int({self.ParseOperator(node.GetChildren()[0], 0)})]"
        for node in operator node.GetChildren():
            lexeme = node.GetLexeme()
            part = None
            if not node.GetChildren():
                if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                    part = str(self.GetVariable(lexeme.itemValue).itemName)
                elif lexeme.itemType in [Language.LexemeTypes.INT NUM,
Language.LexemeTypes.DOUBLE NUM,
                                         Language.LexemeTypes.STRING]:
                    part = str(self.GetLiteral(lexeme.itemValue).itemValue)
            elif node.Type == SyntaxTreNodeTypes.FUNCTION CALL:
                part = self.ParseFunctionCallStatement(node, 0)
            elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                var name = str(self.GetVariable(lexeme.itemValue).itemName)
                part =
f"{var name}[int({self.ParseOperator(node.GetChildren()[0], 0)})]"
            elif IsOperator(operator):
                if lexeme.itemValue == Language.Operators.MINUS and
len(node.GetChildren()) == 1:
                    part = f"{self.ParseOperator(node, 0)}"
                else:
                    part = f"({self.ParseOperator(node, 0)})"
            if part is not None:
                parts.append(part)
            else:
                raise ValueError("Bad operating part")
        return current level + str(operator).join(parts)
    def ParseVariableDeclarationStatement(self, declaration node, level):
            Parses the variable declaration statement to Python.
        declarations = []
        current level = level * '\t'
        for node in declaration node.GetChildren()[1:]:
            lexeme = node.GetLexeme()
```

```
declaration = None
            if not node.GetChildren():
                if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                    variable =
str(self.GetVariable(lexeme.itemValue).itemName)
                    declaration = f"{current level}{variable}=None"
            elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                var name = str(self.GetVariable(lexeme.itemValue).itemName)
                var len = self.ParseOperator(node.GetChildren()[0], 0)
                declaration = f"{current level}{var name}=list(None for in
range({var len}))"
            elif lexeme.itemValue == Language.Operators.EQUAL:
                declaration = f"{self.ParseOperator(node, level)}"
            if declaration is not None:
                declarations.append(declaration)
            else:
                raise ValueError("Unknown declaration")
        return "\n".join(declarations)
    def ParseIfStatement(self, if node, level):
           Parses the if statement to Python.
        current level = level * '\t'
        condition node = if node.GetChildren()[0]
        code node = if node.GetChildren()[1]
        instructions = [self.ParseInstruction(instruction, level + 1) for
instruction in code node.GetChildren()]
        code block = '\n'.join(instructions)
        return f"{current level}if {self.ParseOperator(condition node,
0) }:\n{code block}"
    def ParseFunctionCallStatement(self, call node, level):
           Parses the function call statement to Python.
        current level = level * '\t'
        function name =
self.GetVariable(call node.GetChildren()[0].GetLexeme().itemValue).itemName
        arguments = [self.ParseFunctionArgument(arg)
                     for arg in call node.GetChildren()[1].GetChildren()]
        try:
            return
f"{current level}globals()['{function name}']({','.join(arguments)})"
        except:
            raise ValueError("Bad function call")
    def ParseFunctionArgument(self, argument node):
           Gets the name of the function call argument.
        lexeme = argument node.GetLexeme()
        if not argument node.GetChildren():
            if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
```

```
return str(self.GetVariable(lexeme.itemValue).itemName)
            elif lexeme.itemType in [Language.LexemeTypes.INT NUM,
Language.LexemeTypes.DOUBLE NUM]:
                return str(self.GetLiteral(lexeme.itemValue).itemValue)
            elif lexeme.itemType == Language.LexemeTypes.STRING:
                return repr(self.GetLiteral(lexeme.itemValue).itemValue)
        elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
            var name = str(self.GetVariable(lexeme.itemValue).itemName)
            var len = self.ParseOperator(argument node.GetChildren()[0], 0)
            return f"{var name}[{var len}]"
        elif lexeme.itemType == Language.LexemeTypes.OPERATOR:
            return self.ParseOperator(argument node, 0)
    def ParseWhileStatement(self, while node, level):
           Parses the while cycle statement to Python.
        current level = level * '\t'
        condition node = while node.GetChildren()[0]
        code node = while node.GetChildren()[1]
        instructions = [self.ParseInstruction(instruction, level + 1) for
instruction in code node.GetChildren()]
        code block = '\n'.join(instructions)
        return f"{current level}while {self.ParseOperator(condition node,
0) }: \n{code block}"
    def ParseForStatement(self, for node, level):
            Parses the for cycle statement to Python.
        current level = level * '\t'
        cycle\_level = (level + 1) * '\t'
        condition nodes = for node.GetChildren()
        code node = condition nodes[len(for node.GetChildren()) - 1]
        instructions = [self.ParseInstruction(instruction, level + 1) for
instruction in code node.GetChildren()]
        code block = '\n'.join(instructions)
        if len(condition nodes) == 1 and condition nodes[0]. Type ==
SyntaxTreNodeTypes.CODE BLOCK:
            return f"{current level}while 1:\n{code block}"
        elif len(condition nodes) == 4:
            first condition = None
            if condition nodes[0]. Type == SyntaxTreNodeTypes.DECLARATION:
                first condition =
self.ParseVariableDeclarationStatement(condition nodes[0], level)
            elif condition nodes[0].GetLexeme().itemType ==
Language.LexemeTypes.OPERATOR:
                first condition = self.ParseOperator(condition nodes[0],
level)
            second condition = self.ParseOperator(condition nodes[1], 0)
            third condition = self.ParseInstruction(condition nodes[2],
level)
            return f"{first condition}\n" \
                   f"{current level}while 1:\n" \
                   f"{cycle level}if not {second condition}: break\n" \
```

```
# return f"{current level}while {self.ParseOperator(condition node,
0) }:\n{code block}"
    def ParseDoWhileStatement(self, doWhile node, level):
           Parses the do-while cycle statement to Python.
        current level = level * '\t'
        condition node = doWhile node.GetChildren()[1].GetChildren()[0]
        code node = doWhile node.GetChildren()[0]
        instructions out = [self.ParseInstruction(instruction, level) for
instruction in code node.GetChildren()]
        instructions in = [self.ParseInstruction(instruction, level + 1) for
instruction in code node.GetChildren()]
        code block out = '\n'.join(instructions out)
        code block in = '\n'.join(instructions in)
        return f"{code block out}" \
               f"\n{current level}while {self.ParseOperator(condition node,
0) }: \n{code block in}"
    def ParseUnaryOperatorStatement(self, unary node, level):
            Parses unary operator statements to Python.
        current level = level * '\t'
        var name =
str(self.GetVariable(unary node.GetChildren()[0].GetLexeme().itemValue).itemN
ame)
        if unary node.GetLexeme().itemValue == Language.Operators.INCREMENT:
            return f"{current_level}{var_name}+=1"
        elif unary node.GetLexeme().itemValue ==
Language.Operators.DECREMENT:
            return f"{current level}{var name}-=1"
    def ParseCinStatement(self, cin node, level):
            Parses the input statement to Python.
        current level = level * '\t'
        children = cin node.GetChildren()
        instructions = []
        for node in children:
            lexeme = node.GetLexeme()
            instruction = None
            if not node.GetChildren():
                if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                    arg type =
str(self.GetArgType(self.GetVariable(lexeme.itemValue).itemType))
```

f"{code block}\n" \

f"{current level}{third condition}"

```
instruction = f"{current level}" \
f"{str(self.GetVariable(lexeme.itemValue).itemName)}={arg type}(input())"
            elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                var name = str(self.GetVariable(lexeme.itemValue).itemName)
                var len = self.ParseOperator(node.GetChildren()[0], 0)
                item type =
str(self.GetArgType(self.GetVariable(lexeme.itemValue).itemType[1]))
                instruction = f"{current level}" \
                              f"{var name}[{var len}]={item type}(input())"
            if instruction is not None:
                instructions.append(instruction)
            else:
                raise ValueError("Bad arg to input")
        return '\n'.join(instructions)
    def ParseCoutStatement(self, cout node, level):
            Parses the output statement to Python.
        current level = level * '\t'
        children = cout node.GetChildren()
        messages = []
        for node in children:
            lexeme = node.GetLexeme()
            message = None
            if not node.GetChildren():
                if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                    message =
str(self.GetVariable(lexeme.itemValue).itemName)
                elif lexeme.itemType in [Language.LexemeTypes.INT NUM,
Language.LexemeTypes.DOUBLE NUM]:
                    message =
str(self.GetLiteral(lexeme.itemValue).itemValue)
                elif lexeme.itemType == Language.LexemeTypes.STRING:
                    message =
repr(self.GetLiteral(lexeme.itemValue).itemValue)
                elif lexeme.itemValue == Language.KeyWords.ENDL:
                    message = repr("\n")
            elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                var name = str(self.GetVariable(lexeme.itemValue).itemName)
                var len = self.ParseOperator(node.GetChildren()[0], 0)
                message = f"{var name}[{var len}]"
            if message is not None:
                messages.append(message)
            else:
                raise ValueError("Bad arg to print")
        return f"{current level}print({','.join(messages)}, end={repr('')})"
    def ParseReturnExitStatement(self, return exit node, level):
            Parses the return and exit statements to Python.
        current level = level * '\t'
```

```
children = return exit node.GetChildren()
        return arg = None
        if children:
            node = children[0]
            lexeme = node.GetLexeme()
            if not node.GetChildren():
                if lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                    return arg =
str(self.GetVariable(lexeme.itemValue).itemName)
                elif lexeme.itemType in [Language.LexemeTypes.INT NUM,
Language.LexemeTypes.DOUBLE NUM]:
                    return arg =
str(self.GetLiteral(lexeme.itemValue).itemValue)
                elif lexeme.itemType == Language.LexemeTypes.STRING:
                    return arg =
repr(self.GetLiteral(lexeme.itemValue).itemValue)
            elif lexeme.itemType == Language.LexemeTypes.IDENTIFIER:
                var name = str(self.GetVariable(lexeme.itemValue).itemName)
                var len = self.ParseOperator(node.GetChildren()[0], 0)
                return arg = f"{var name}[{var len}]"
            if return arg is None:
                raise ValueError("Bad return argument")
            else:
                if return exit node.GetLexeme().itemValue ==
Language.KeyWords.RETURN:
                    return f"{current level}return {return arg}"
                elif return exit node.GetLexeme().itemValue ==
Language.KeyWords.EXIT:
                    return f"{current level}quit({return arg})"
            if return exit node.GetLexeme().itemValue ==
Language.KeyWords.RETURN:
                return f"{current level}return"
            elif return exit node.GetLexeme().itemValue ==
Language.KeyWords.EXIT:
                return f"{current level}quit()"
```